electronics

Review

Single-Image Super-Resolution Challenges: A Brief Review

Shutong Ye [1], Shengyu Zhao [1], Yaocong Hu [2] and Chao Xie [1,3,*]

1 College of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing 210037, China
2 School of Electrical Engineering, Anhui Polytechnic University, Wuhu 241000, China
3 College of Landscape Architecture, Nanjing Forestry University, Nanjing 210037, China
* Correspondence: chaoxie@njfu.edu.cn

Abstract: Single-image super-resolution (SISR) is an important task in image processing, aiming to achieve enhanced image resolution. With the development of deep learning, SISR based on convolutional neural networks has also gained great progress, but as the network deepens and the task of SISR becomes more complex, SISR networks become difficult to train, which hinders SISR from achieving greater success. Therefore, to further promote SISR, many challenges have emerged in recent years. In this review, we briefly review the SISR challenges organized from 2017 to 2022 and focus on the in-depth classification of these challenges, the datasets employed, the evaluation methods used, and the powerful network architectures proposed or accepted by the winners. First, depending on the tasks of the challenges, the SISR challenges can be broadly classified into four categories: classic SISR, efficient SISR, perceptual extreme SISR, and real-world SISR. Second, we introduce the datasets commonly used in the challenges in recent years and describe their characteristics. Third, we present the image evaluation methods commonly used in SISR challenges in recent years. Fourth, we introduce the network architectures used by the winners, mainly to explore in depth where the advantages of their network architectures lie and to compare the results of previous years' winners. Finally, we summarize the methods that have been widely used in SISR in recent years and suggest several possible promising directions for future SISR.

Keywords: single-image super-resolution; single-image super-resolution challenges; deep learning; deep networks

Citation: Ye, S.; Zhao, S.; Hu, Y.; Xie, C. Single-Image Super-Resolution Challenges: A Brief Review. *Electronics* **2023**, *12*, 2975. https://doi.org/10.3390/electronics12132975

Academic Editors: Beiwen Li, Vignesh Suresh and Yi Zheng

Received: 17 May 2023
Revised: 25 June 2023
Accepted: 4 July 2023
Published: 6 July 2023

1. Introduction

Single-image super-resolution is an important task in image processing, aiming to reconstruct high-resolution images from low-resolution images and optimize both details and textures to improve the quality of visual perception. It is currently used in a wide range of real-life scenarios [1–4], including security surveillance [5–7], remote sensing [8–10], medical imaging [11–13], etc., while contributing to other advanced computer vision tasks [14–19], and is therefore of wide interest to academia and industry [20–24].

With the rapid development of deep learning [25–30], deep-learning-based SISR models have achieved state-of-the-art performance on various benchmarks, but SISR remains a challenging problem as a severely discomforting computer vision problem. This discomfort can become more severe as the scale factor changes, so there are still many aspects of SISR that need to be improved.

To facilitate the development of SISR, challenges regarding image super-resolution have emerged. Among them, NTIRE, PIRM, and AIM are the three most popular challenges. In this paper, we overview the recent progress of deep-learning-based SISR in addressing its top challenge. Although there have been some previous surveys on SISR [31–36], our survey differs from them in that we focus on the performance and progress of SISR techniques that address their top challenges. Unlike earlier works that mostly investigated traditional SISR algorithms or focused on a particular class of SISR techniques, this survey systematically and comprehensively reviews the development of SISR as its top challenge during 2017–2022.

NTIRE: The New Trends in Image Recovery and Enhancement (NTIRE) challenge was combined with CVPR [37]. For the single-image super-resolution challenge, the challenge tasks include efficient super-resolution, extreme super-resolution, real-world super-resolution, and classic super-resolution, intending to reconstruct a degraded and resulting low-resolution image into a new high-resolution image at a target multiple, and the challenge promotes the development of SR research in the ideal case or the real-world case.

PIRM: The Perceptual Image Recovery and Manipulation (PIRM) challenge was held in conjunction with ECCV and includes multiple tasks [38]. This challenge focuses on generating high-resolution images with both accuracy and perceptual quality. It is well known that when you choose to generate high-resolution images with higher accuracy, you tend to obtain poor visual perception, and when you choose to generate high-resolution images with higher perceived quality, the image quality is often not good enough.

AIM: The Advances in Image Manipulation (AIM) challenge was combined with ICCV [39,40]. The AIM challenge includes the following main tasks: to train SISR models that can be applied to real-world scenarios, to improve the efficiency of SISR models, to increase the speed of the models, and to reduce the memory needed to run them given a benchmark, etc.

This paper mainly reviews the content of the challenges and the superior methods on single-image super-resolution in NTIRE, PIRM, and AIM during 2017–2022. The rest of this paper is organized as follows: Section 2 presents the datasets used in the above challenge; Section 3 presents the various IQAs proposed and used in the challenge; Section 4 presents the models that won the above challenge during 2017–2022, focusing on the deep feature extraction part; Section 5 concludes and discusses the possible future directions of SISR.

2. Background

Among the SISR tasks, we can express the degradation process of high-resolution images to low-resolution images using the following formula:

$$y = \varphi(x, \theta_\eta) \tag{1}$$

where y denotes a low-resolution image, x denotes a high-resolution image, φ is a function representing the degradation process, and θ_η is various parameters in the degradation process, including noise and downscaling kernels. And the SISR task is to try to predict and reconstruct a high-resolution image $\hat{x}$ from the degraded obtained low-resolution image, and the process can be expressed as follows:

$$\hat{x} = \varphi^{-1}(y, \theta_\varsigma) \tag{2}$$

where $\hat{x}$ denotes the reconstructed high-resolution image, y denotes the input low-resolution image, φ^{-1} denotes the function solved backward from the degradation process, and θ_ς is the various parameters in the function solved backward. The image degradation process in SISR tasks is often unknown and complex and is affected by various factors, such as noise, blur, compression, and artifacts, so the most challenging task in SISR tasks is how to construct the inverse solution function φ^{-1}. In the field of SISR, most researchers have modeled the degenerate function φ in Equation (1) as follows:

$$y = (x \otimes k) \downarrow_s + n \tag{3}$$

where y denotes a low-resolution image, x denotes a high-resolution image, $\otimes$ denotes a convolution operation, k denotes a blurring kernel that makes the image blurred, $\downarrow_s$ denotes a downscaling operation that reduces the size of the image by a factor of s, and n denotes an additive Gaussian white noise with kernel width σ.

3. Dataset

To compare the strengths and weaknesses of the SISR model, it is necessary to train as well as test the validation on the same dataset. The datasets used until 2017 are Train 91 [41] proposed by Yang et al. and Set5 [42], Set14 [43], BSD100 [44], and Urban100 [45] proposed and merged by Timofte et al. [46].

With the development of SISR networks, the size of the previously proposed dataset is not sufficient for training complex neural networks, and the training of SISR networks requires more a priori information, so the size of the dataset gradually increases.

First presented at NTIRE 2017, the DIV2K [47] dataset features 1000 images collected from the Internet and covers a variety of content, including people, environments, animals, and more. Each image in this dataset has 2K pixels, i.e., they have 2K pixels in at least one axis (horizontal or vertical) direction with a much higher resolution than the images in the dataset presented above.

The challenging task of NTIRE 2019 was to achieve single-image super-resolution in the real world. At that time, most of the LR images in the dataset were HR images obtained via simple bicubic downscaling, while the image degradation in the real world was much more complex than that, so the SISR method at that time did not perform well on the real-world images. The dataset used in NTIRE 2019 is RealSR [48], proposed by J. Cai et al. This dataset was captured using a digital camera, and an image alignment algorithm was developed to gradually align image pairs at different resolutions to obtain LR-HR image pairs of the same scene by adjusting the focal length. In addition, for the 2020 AIM Real-World Super-Resolution Challenge, Wei et al. proposed the DRealSR [49] dataset, which has more numbers and diversity than the RealSR, and the DPED [50] dataset, which consists of three different mobile phones and one high-end camera, was used for the challenge. The dataset consists of real photos taken on three different cell phones and a high-end camera.

Due to the resolution limitation of the dataset, scaling of larger factors was difficult to achieve with the then-available datasets, so the DIV8K [51] dataset was proposed at the 2019 AIM Extreme Super-Resolution Challenge, which is suitable for scaling factors of 32 and above. The dataset has 1504 high-resolution images, of which the validation set and the test set have one hundred images each. The horizontal pixel resolution of the images in the validation set, test set, and part of the training set is not less than 7680, and the horizontal resolution of the remaining images in the training set is not less than 5760. In Table 1, we list a number of datasets commonly used in the SISR challenges. In Figure 1, We show a selection of images from a commonly used dataset.

Table 1. Image representation of the SISR challenge datasets 2017–2022.

Dataset	Amount	Format	Short Description
Train 91	91	PNG	Images for training, including a car, flower, fruit, etc.
Set5	5	PNG	Images for testing, including a baby, bird, butterfly, head, and woman.
Set14	14	PNG	Images for testing, including humans, animals, insects, etc.
BSD100	100	PNG	Images for testing, including animals, buildings, food, etc.
Urabn100	100	PNG	Images for testing, including a city, urban, structure, etc.
DIV2K	1000	PNG	Each image in this dataset has 2K pixels, including the environment, flora, fauna, handmade object, etc.
RealSR	595	PNG	The dataset was built via two cameras (Cannon 5D3 and Nikon D810) and used an image alignment algorithm to obtain LR-HR pairs for the real-world SISR challenge.
DRealSR	2507	PNG	Compared to RealSR, it has more diversity and more data volume.
DPED	6000	PNG	The authors used three cell phones and a DSLR to photograph an object simultaneously to form a new database of 6000 photographs for this study.
DIV8K	1504	PNG	Images are suitable for scaling factors of 32 and above.

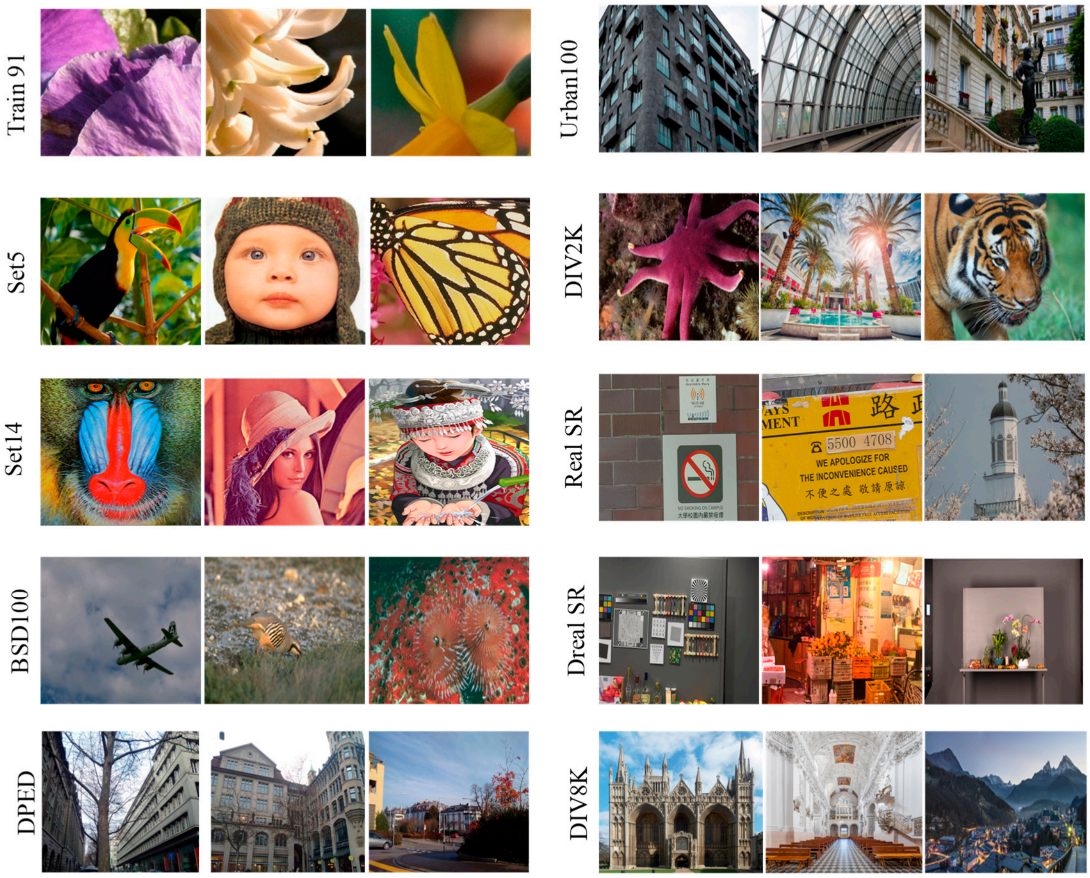


Figure 1. Image representation of the SISR challenge datasets 2017–2022.

4. Evaluation Method

Typically, image quality is assessed using both subjective human perception methods (i.e., whether the image looks realistic) and objective methods. SISR aims to generate images that match human perception and are of high image quality. Since subjective human perception methods take a lot of time to evaluate, the prevailing method is the objective method. Since objective methods do not reflect human perception of images, the results obtained using subjective and objective methods sometimes differ significantly, and we next describe the subjective and objective methods used in the SISR challenge in 2017–2022.

4.1. Peak Signal-to-Noise Ratio (PSNR)/Structural Similarity Index (SSIM)

Given a high-resolution image as opposed to low-resolution images I with N pixels and a super-resolution image $\hat{I}$, L is typically 255, and PSNR [52] is defined based on MSE.

MSE is defined as follows:

$$\text{MSE} = \frac{1}{N}\left\| I - \hat{I} \right\|^2 \tag{4}$$

PSNR is defined as follows:

$$\text{PSNR} = 10log_{10}\left(\frac{L^2}{\text{MSE}}\right) \tag{5}$$

SSIM [52] is defined as follows:

$$SSIM(I,\hat{I}) = \frac{(2\mu_I\mu_{\hat{I}} + C_1)(\sigma_{I\hat{I}} + C_2)}{\left(\mu_I^2 + \mu_{\hat{I}}^2 + C_1\right)\left(\sigma_I^2 + \sigma_{\hat{I}}^2 + C_2\right)} \tag{6}$$

where μ_I and σ_I^2 are the mean and variance of I, $\sigma_{I\hat{I}}$ is the covariance between I and $\hat{I}$, and C_1 and C_2 are the constant terms.

4.2. Perception Index (PI)

The previous IQA can only reflect the quality of the picture, which is difficult to reflect the effect of human visual perception of the picture. The PI is proposed in the 2018 PIRM Challenge on Perceptual Image Super-resolution to reflect the perceived quality of the picture.

The No-Reference Quality Metric (NRQM) [53] is a learning-based no-reference metric that trains a regression network for evaluating the perceptual quality of SR images by learning a large number of SR images and the corresponding perceptual scores.

Natural Image Quality Evaluator (NIQE) [54] is a natural scene statistic (NSS) model based on which the quality of the test image is expressed as the distance between the multivariate Gaussian (MVG) fit of the NSS features extracted from the test image and the MVG model of the perceptual quality features extracted from the natural image. PI uses reference-free image quality assessment methods such as NRQM and NIQE to achieve the following:

$$\mathrm{PI} = \frac{1}{2}((10 - \mathrm{NRQM}) + \mathrm{NIQE}) \tag{7}$$

4.3. Learned Perceptual Image Patch Similarity (LPIPS)

LPIPS [55] is used to measure the difference between two images and is more consistent with human perception than the traditional SSIM and PSNR.

$$d(x,x_0) = \sum_l \frac{1}{H_l W_l} \sum_{h,w} \left\| w_l \odot \left(\hat{y}_{hw}^l - \hat{y}_{0hw}^{\,l}\right)\right\|_2^2 \tag{8}$$

where d denotes the distance from x_0 to x. The feature stack is extracted from the l layer and unit-normalized in the channel dimension. The vector W_l is used to deflate the number of activated channels, and finally, the L_2 distance is calculated. Finally, it is averaged over the space and summed by channel.

4.4. Mean Opinion Score (MOS)/Mean Opinion Rank (MOR)

MOS refers to the scoring of the generated image against a relative reference image with six ratings.

$$\mathrm{MOS} = \sum_{x\in[0,1,2,3,4,5]} x \cdot p(x) \tag{9}$$

MOR means that the study participants are asked to rank the images obtained using the different methods without seeing the reference image. In addition, the IQA-Rank was calculated using the average of four evaluation methods: NIQE, Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [56], Perception-based Image Quality Evaluator (PIQE) [57], and NRQM. BRISQUE means extracting the mean subtracted contrast normalized (MSCN) features from the image and fitting them to an asymmetric generalized Gaussian distribution (AGGD), extracting the fitted Gaussian distribution features and inputting them to a vector machine (SVM) to predict the result of the image quality evaluation. PIQE is more concerned with extracting local features and predicting the overall image quality score from the local image quality score. Similar to BRISQUE, PIQE also calculates MSCN coefficients first and then calculates the quality fraction of the whole picture according to the formula based on distortion, etc. Then, the IQA-Rank of each method and the ranking given by the participants were used to calculate the average to obtain the MOR.

4.5. Parameters Used to Measure Efficiency

To improve the operational efficiency of the network, several parameters used to measure efficiencies, such as runtime, parameter calculation, FLOPs, activation, and GPU memory usage, are used in a series of efficient super-resolution challenges. In Table 2, we show the subjective and objective methods used in the SISR challenge in 2017–2022.

Table 2. The subjective and objective methods used in the SISR challenge in 2017–2022.

Evaluation Method	Full-/Non-Reference	Short Description
Peak Signal-to-Noise Ratio (PSNR)	Full-Reference	The image quality reference value between the maximum signal and the background noise is calculated to measure whether the image is distorted or not. Higher PSNR values indicate higher quality of the generated images.
Structural Similarity Index (SSIM)	Full-Reference	Calculate whether the structure of two images is similar from the point of view of brightness, contrast, and image structuring. The larger the SSIM, the more similar the images are.
Perception Index (PI)	Non-Reference	It is used to calculate the perceived quality of the image, and often the lower the value, the better the perceived quality of the image.
Learned Perceptual Image Patch Similarity (LPIPS)	Full-Reference	The perceptual similarity is calculated, which is more in line with human perception than traditional methods (PSNR and SSIM). The lower value indicates that the two images are more similar.
Mean Opinion Score (MOS)	Non-Reference	The perceived quality of the images is evaluated via human ratings.
Mean Opinion Rank (MOR)	Non-Reference	Similar to MOS, the perceived quality of the images is evaluated via human ratings of the images.

5. Superior Method

Depending on the task of the challenge, it can be broadly classified into four categories: classic SISR, efficient SISR, perceptual extreme SISR, and real-world SISR. In this section, the network architecture of the winning approach in the 2017–2022 SISR challenge will be presented.

5.1. Classic SISR

Classic SISR refers to the reconstruction of LR images obtained by bicubic downsampling or unknown degradation into images with magnification factors of $\times 2$, $\times 3$, $\times 4$. The classic SISR challenge has two tracks: one track is to obtain the LR image corresponding to each HR using the classic bicubic downsampling and degradation factor; the other track is to obtain the LR image using an unknown degradation. The goal of both tracks is to reconstruct the original HR image from LR separately [37,47,58,59]. In Table 3, we show the classic SISR challenge track and winner for the period 2017–2022. In Table 4, we show the results of the classic SISR challenge winners with amplification factors of 2, 3, and 4 for 2017–2022 and show the winner with a factor of 8 in Table 5.

Table 3. Classic SISR challenge track.

Challenge	Track	Winner
NTIRE 2017	Track 1: Degradation is achieved using bicubic downscaling with degradation factors ×2, ×3, ×4 Track 2: Degradation is achieved using an unknown method with degradation factors ×2, ×3, ×4	EDSR
NTIRE 2018	Track 1: Degradation is achieved using bicubic downscaling with a degradation factor of ×8 Track 2: Degradation is achieved using an unknown method with a degradation factor of ×4. Track 3: Similar to Track 2, except that the degradation is more complex, with a degradation factor of ×4. Track 4: Similar to Track 2 and 3, the degradation factor is ×4, and the degradation mode is different between images, and four LR images are generated for each HR image.	DBPN WDSR
AIM 2022	Implement a ×4 super-resolution for JPEG images compressed using python code with a quality factor of 10.	TCIR

Table 4. Results of the classic SISR challenge winners (×2, ×3, and ×4) in 2017–2022.

Challenge Category	SISR Networks	×2		×3		×4	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
classic	EDSR	34.93	0.948	31.13	0.889	29.09	0.837
blind	EDSR	34.00	0.934	30.78	0.881	28.77	0.826

Table 5. Results of the classic SISR challenge winners (×8) in 2017–2022.

Challenge Category	SISR Networks	×8	
		PSNR	SSIM
classic	DBPN	25.455	0.7088
blind	WDSR (Mild)	23.631	0.6316
	WDSR (Difficult)	22.329	0.5721
	WDSR (Wild)	23.080	0.6038

5.1.1. EDSR (Winner of NTIRE 2017)

The EDSR [60] proposed by the SNU CVLab team won in both tracks. The SNU CVLab team has made a series of improvements to the EDSR based on SRResNET [61], and the specific network structure is shown in Figure 2. They removed the Batch Normalization (BN) from the residual module because the BN [62] layer would have eliminated the flexibility of the network due to the normalization feature, and this operation was effective in improving the PSNR. The training process becomes unstable as the depth of the network increases. To solve this problem, EDSR uses residual scaling. A residual scaling layer (scaling using constant multiplication) is also added after the second convolution, which is experimentally found to effectively stabilize the learning process when C = 0.1. The model consists of 36 such residual modules. In Figure 3 we compare the original residual block with the residual block in EDSR. The EDSR only has upsampling modules that differ by the scale factor. The EDSR architecture further optimizes SRResNET by removing the BN layer from the SISR network to improve the network performance, allowing one to train a larger model under limited conditions, and also giving rise to the exploration of batch normalization layers in SISR networks.

Figure 2. The overall architecture of the EDSR network.

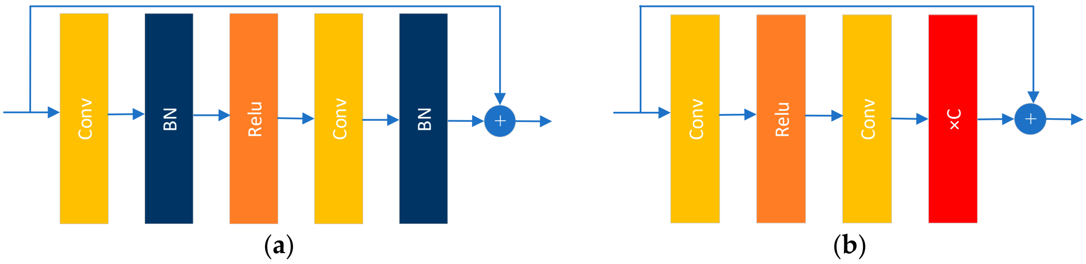


Figure 3. Comparison of the original residual block (**a**) and the residual block in EDSR (**b**).

5.1.2. DBPN/WDSR (Winner of NTIRE 2018)

The DBPN [63] proposed by Muhammad Haris et al. achieved superiority on track 1, and the specific network structure is shown in Figure 4. DBPN is a kind of back-projection network, the previous networks are more feed-forward to predict the results of SR, and each layer is basically based on the results of the previous layer to obtain. We show the upper and lower projection units in DBPN in Figure 5. DBPN connects the features of up- and downsampling together. Each instance of upsampling or downsampling reconstruction uses all of the previous LR or HR image features. Finally, splicing the depth features of all the HR images were obtained from upsampling to reconstruct the final HR image. Unlike previous feed-forward networks, DBPN proposes an iterative mapping projection network that fully exploits the relationship between low-resolution images and high-resolution images, uses the error between the upper and lower projections to guide the reconstruction of images, and stitches the feature maps of all the high-resolution images obtained via upsampling to reconstruct high-resolution images. DBPN also achieves a super-resolution network with a large magnification factor.

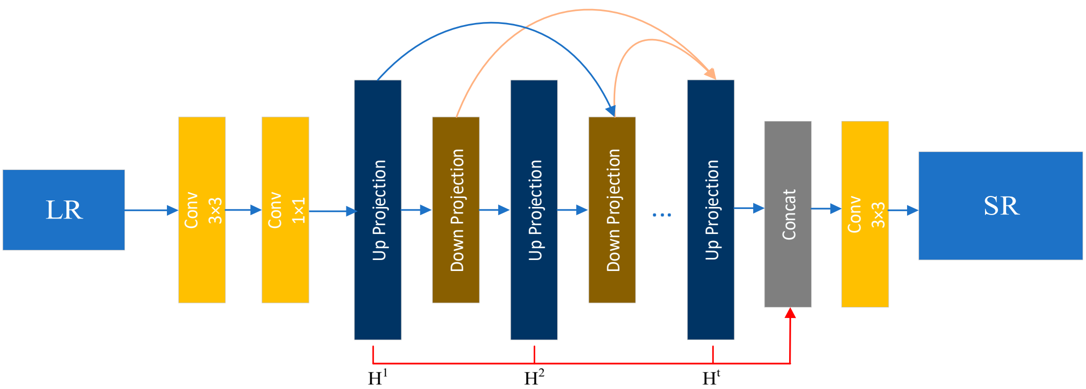


Figure 4. The overall structure of the DBPN network.

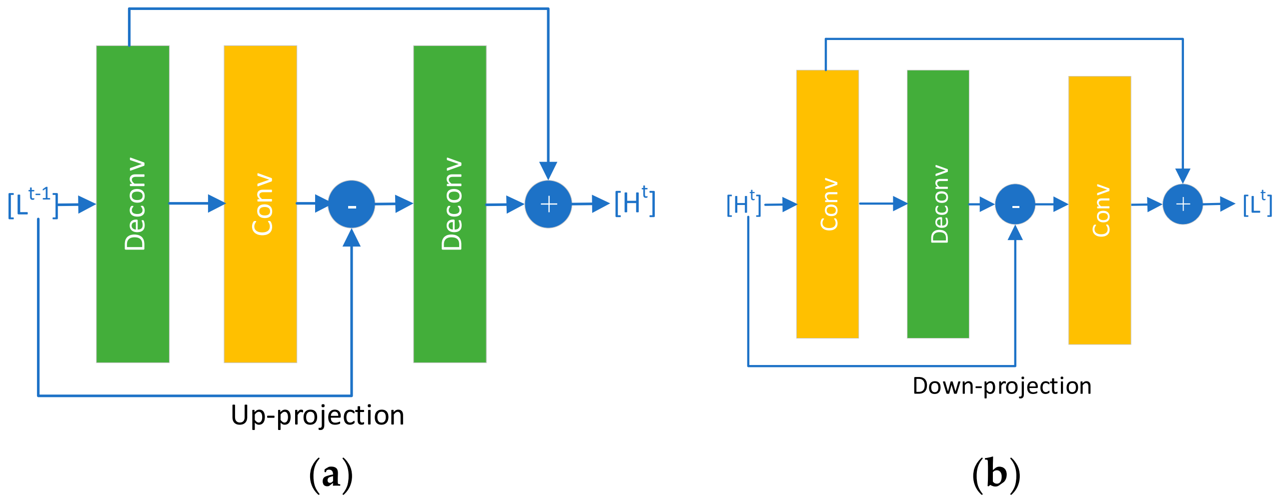


Figure 5. Upper projection units (**a**) and lower projection (**b**) units in DBPN.

The WDSR [64] proposed by Yu et al. achieved superiority on track 2, and the specific network structure is shown in Figure 6. WDSR has improved the Residual block in EDSR by increasing the number of channels of the feature map before the ReLU function, which can activate the network better and obtain better performance without increasing the computational overhead; In addition, a large convolutional kernel after the ReLU function is split into two small convolutional kernels, which can effectively reduce the number of computations while ensuring the same perceptual field; in addition, WDSR replaces the BN layer with the Weight Normalization (WN) [65] layer to increase the training speed and speed up the convergence of the neural network.

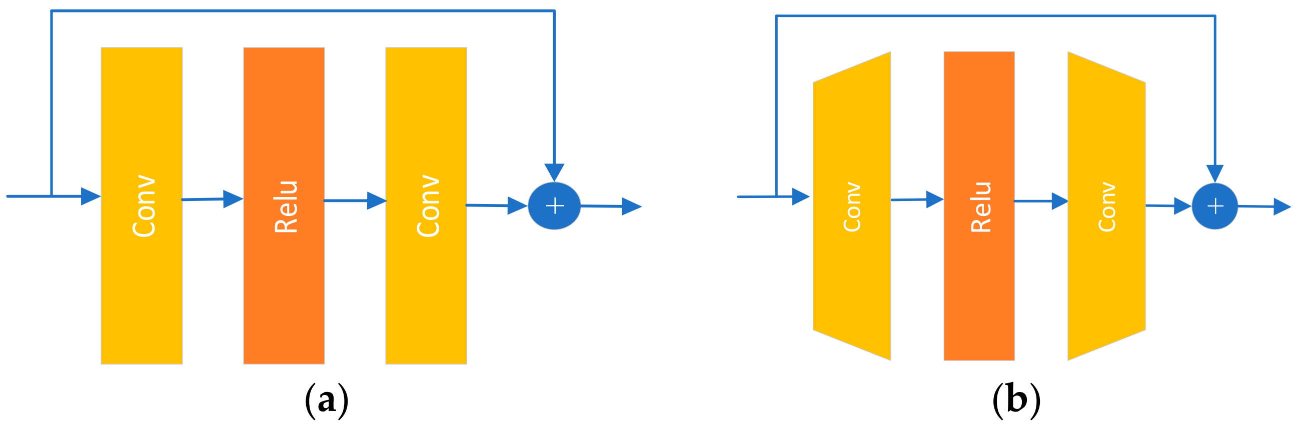


Figure 6. Residual block in EDSR (**a**) and residual block in WDSR (**b**).

In the overall network architecture, WDSR removes the redundant convolutional layers from the EDSR and does not insert convolutional blocks after the upsampling layer. We show that in Figure 7. This operation can effectively improve the operation speed and reconstruction effect of the network.

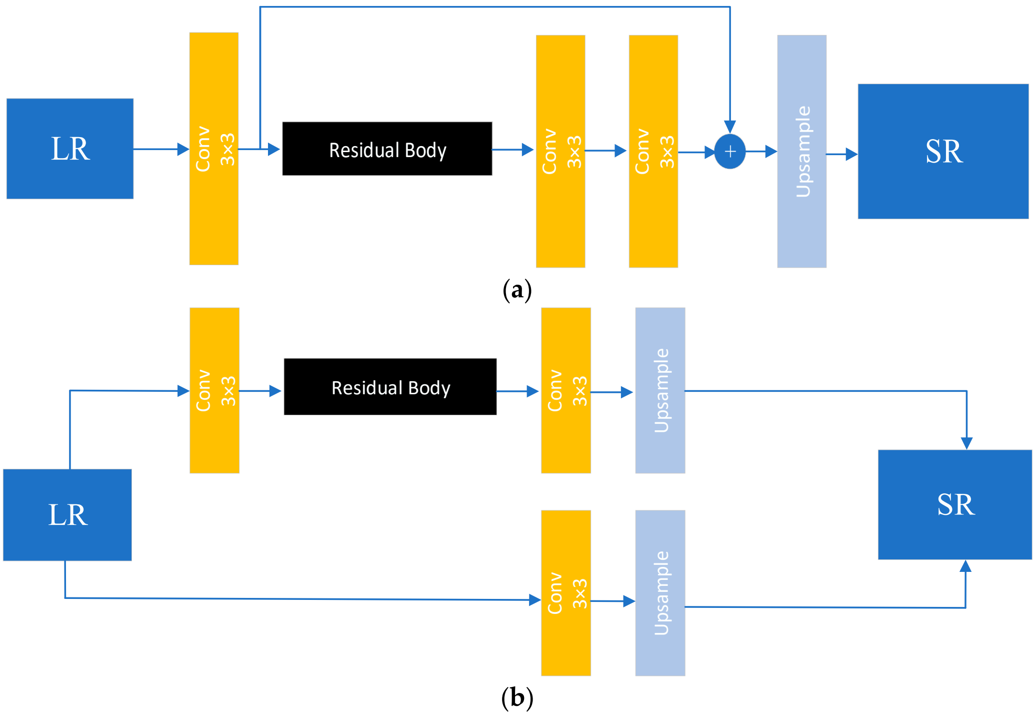


Figure 7. Simplified results of WDSR (**b**) compared to EDSR (**a**).

5.1.3. TCIR (Winner of AIM 2022)

Compression plays an important role in the efficient transmission of images on broadband-limited Internet, but compression can lead to image artifacts and degrade image quality. Therefore, AIM 2022 proposed a super-resolution challenge for compressed images using the DIV2K dataset, and the TCIR proposed by the VUE Team won the year, with the specific network structure shown in Figure 8.

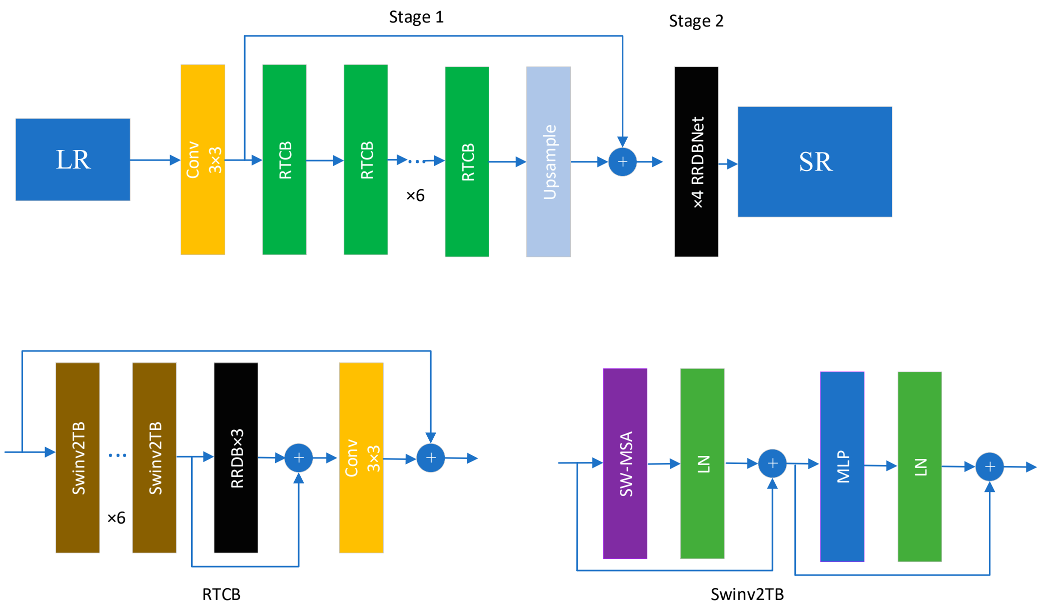


Figure 8. TCIR's overall network architecture.

They divided the network into two stages: the first using a hybrid network of Transformer and CNN to remove artifacts and the second using a modified RRDBNet to achieve $\times 4$ super-resolution.

The improvements made by the team to SwinIR [66] are outlined below. Firstly, they downsample the image with a convolution of step 2 to shrink the image by a factor of two. Since the image itself is compressed with a quality factor of 10, this operation does not affect the performance of TCIR, saves GPU memory, and accelerates the model. And then they use the new Swinv2 transformer module to replace the STL module in SwinIR to greatly improve the performance of the network. Thirdly, they added three RRDB modules to the RTCB (the basic module of the network) in TCIR, which can take advantage of both CNN and the Transformer. The network combines CNN and the Transformer together and achieves excellent results, proving that the combination of CNN and the Transformer has a good development prospect.

5.2. Efficient SISR

The goal of the Efficient SISR Challenge is to increase the efficiency of super-resolution networks with amplification factors of $\times 2$, $\times 3$, and $\times 4$ as much as possible. Factors that affect the efficiency of SISR networks include runtime, number of parameters, FLOPS, activation, and memory consumption. Therefore, the efficiency of the SISR network is evaluated from different aspects. The Efficient SISR Challenge is a SISR challenge that optimizes other metrics as much as possible while limiting one parameter, aiming to improve the operational efficiency of SISR networks and advance the lightweight and efficiency of networks [39,67,68]. In Table 6, we show the efficient SISR challenge track and winner for the period of 2017–2022, and, in Table 7, we show the results of efficient SISR challenge winners for 2017–2022.

Table 6. Efficient SISR Challenge track in 2017–2022.

Challenge	Track	Winner
AIM 2019	Track 1: Degradation is achieved using bicubic downscaling with degradation factors $\times 2$, $\times 3$, $\times 4$ Track 2: Degradation is achieved using an unknown method with degradation factors $\times 2$, $\times 3$, $\times 4$ Track 3: Fidelity is used to design networks with high fidelity under the premise of guaranteeing the PSNR and running time of MSRResNet.	IMDN BaiDu-NAS
AIM 2020	The goal of this challenge is to design a network that reduces one or more aspects, such as runtime, parameters, FLOP, activation, and memory consumption while maintaining at least the PSNR of MSRResNet.	RFDN
NTIRE 2022	Main Track: Designing networks with short run times. Sub Track 1: Designing networks with few model parameters and FLOPS. Sub Track 2: Combines runtime, parameters, FLOPs, activation, and memory consumption.	RLFN BSRN EFDN

Table 7. Results of the Efficient SISR Challenge winners. 'Params' denotes the total number of parameters in 2017–2022. 'FLOPs' is the abbreviation for floating point operations. 'Acts' measures the number of elements of all outputs of convolutional layers. 'GPU Mem' represents maximum GPU memory consumption.

SISR Networks	PSNR [dB]	Ave. Time [ms]	Params [M]	FLOPs [G]	Acts [M]	GPU Mem. [M]
IMDN	28.78	50.86	0.893	58.63	154.14	120
BaiDu NAS	28.84	-	1.461	-	-	-
RFDN	28.75	41.97	0.433	27.10	112.03	200
RLFN	28.72	27.11	0.317	19.70	80.05	377.91
BSRN	28.69	140.47	0.156	9.50	65.76	729.94
EFDN	28.71	29.97	0.272	16.86	79.59	575.99
MSRResNet	28.70	-	1.517	166.36	292.55	610

5.2.1. IMDN (Winner of AIM 2019)

AIM 2019 presents a challenge of image super-resolution performed under constraints, which is based on MSRResNet [69] as a baseline with three tracks. The aim is to optimize the remaining one parameter under two of the conditions limiting the number of parameters, runtime, and PSNR, respectively.

The winner of the year was the IMDN (Information Distillation Network) [70] proposed by the Rainbow Team, the specific network structure of which is shown in Figure 9. The main idea is to use the IMDB module to replace the 16 residual modules in MSRResNet, where the IMDB module is shown in Figure 10. This module can divide the intermediate features into two parts by channels, one part is retained, and the other part is further processed via a 3 × 3 convolutional layer, and a 1 × 1 convolution is used to combine them at the end. This operation can effectively preserve information and greatly improve the performance of the SISR network with only a small increase in parameters. The final upsampling module simply employs a sub-pixel convolution to preserve as many parameters as possible. IMDB uses a Contrast-aware Channel Attention layer (CCA) to enhance image details and improve the accuracy of SISR. Due to the split channel operation in extracting features, the number of input channels is reduced, and an excellent balance between the number of parameters, running time, and PSNR at runtime is achieved. The information distillation model proposed via IMDN is one of the most advanced methods for lightweight networks, which effectively guides the development of lightweight networks.

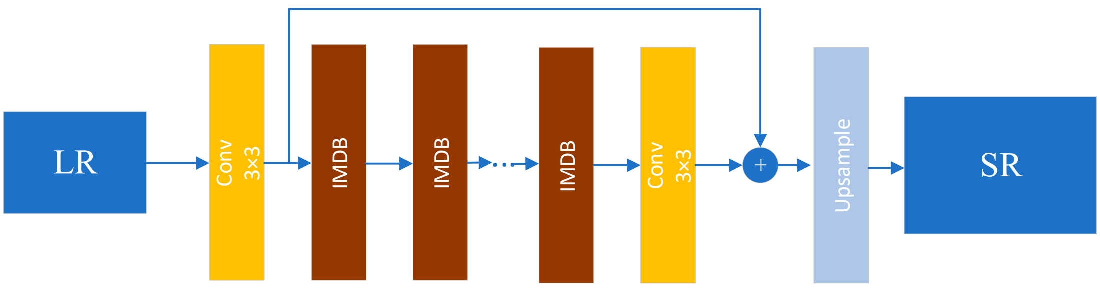


Figure 9. Overall structure of IMDN.

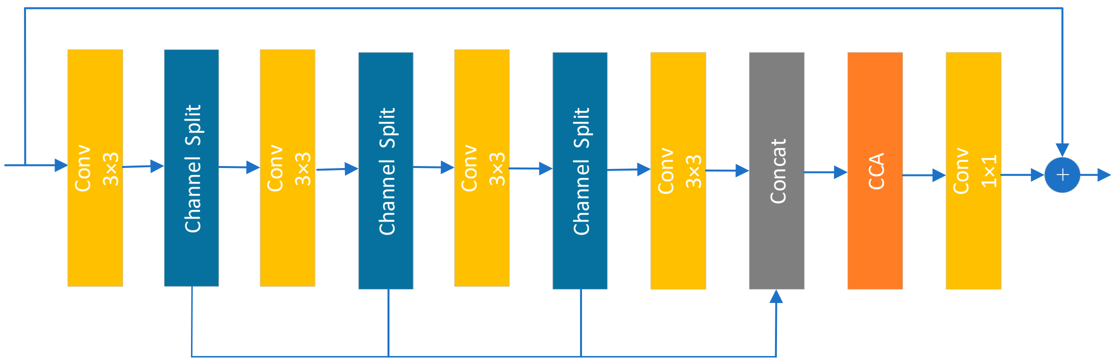


Figure 10. IMDB module.

5.2.2. RFDN (Winner of AIM 2020)

The 2020 AIM presents an efficient single-image super-resolution challenge with a magnification factor of ×4. The goal is to design a network that reduces one or more aspects of runtime parameter computation, flops, activation, and memory consumption while guaranteeing the PSNR of MSRResNet.

The RFDN [71] proposed by the NJU MCG team achieved superiority, and the specific network structure is shown in Figure 11. The team proposes the FDC module to make the network lighter and more accurate and proposes the SRB module to enable the network to harvest the most from the residual learning.

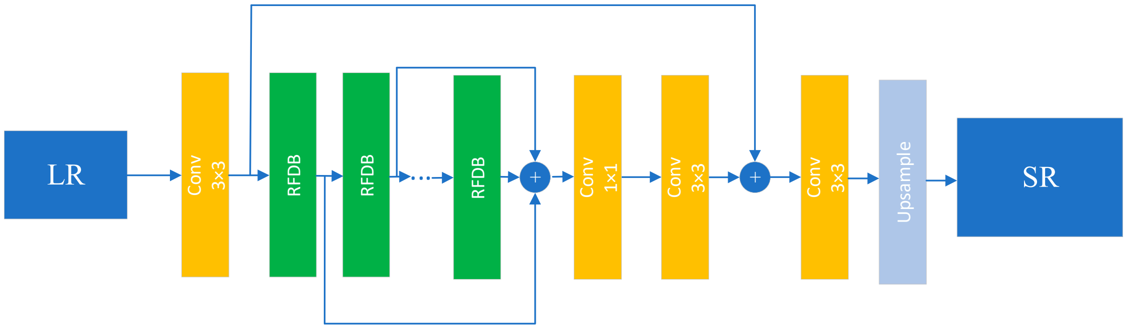
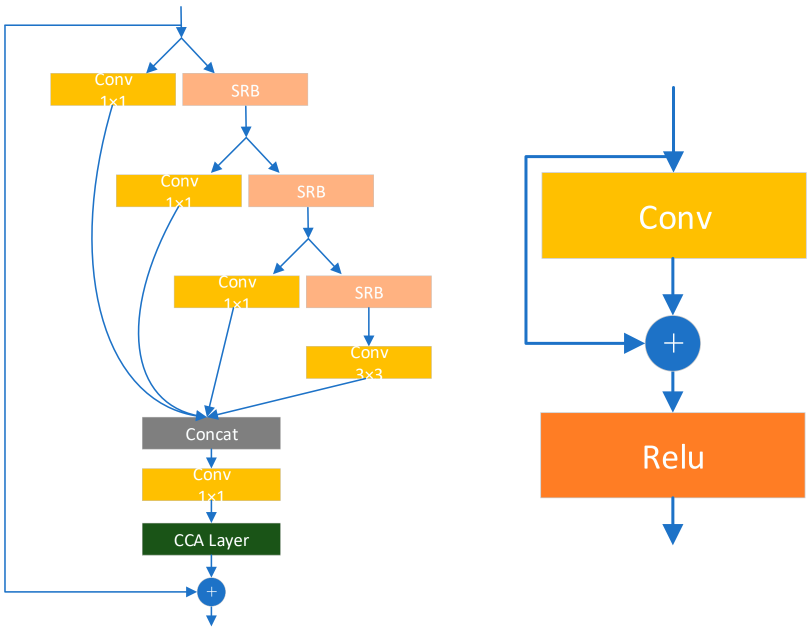


Figure 11. RFDN overall network structure.

The NJU MCG team found that the feature extraction operation is implemented with a 3×3 convolution, but similar to other CNN models, it is more efficient to use a 1×1 convolution for channel separation. To ensure the spatial context and better refine the features, the 3×3 convolution is still used on the right convolutional theme, which is the FDC block proposed by the team.

The team also introduced a smaller range of residual learning in the network by designing a shallow residual block (SRB), which consists of a 3×3 convolution, a connection, and an activation unit, as shown in Figure 12. This block can benefit from residual learning without introducing any additional parameters; the residual linking in IMDB is too coarse to take advantage of residual linking. In contrast, the SRB block allows a lightweight network to take advantage of residual learning as well. In addition, the authors believe that using spatial attention in a shallow SR model is more effective than using channel attention and, therefore, replaced the CCA layer in the RFDB with the ESA layer in the network model that participated in the competition.



(**a**) (**b**)

Figure 12. RFDB (**a**) and SRB (**b**) in RFDN.

5.2.3. RLFN (Winner of NTIRE 2022)

In 2022, NTIRE proposed the Single-Image Efficient Super-Resolution Challenge, which has three tracks [67]: a main track for runtime, a sub-track for model complexity, and a sub-track for overall performance. The RLFN [72] proposed by ByteESR won in the main track, and the specific network structure is shown in Figure 13.

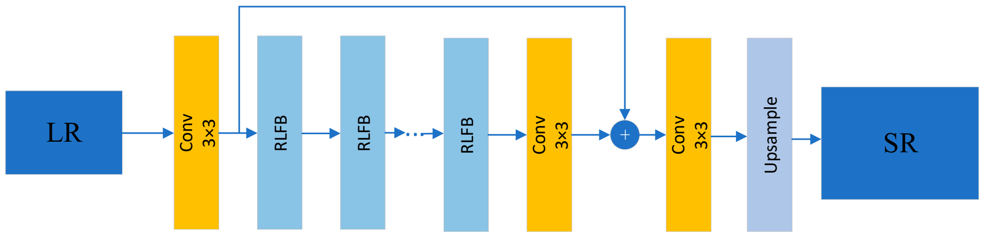


Figure 13. The overall network architecture of RLFN.

The team rethought RFDB and proposed the RLFB module as the basic module of their network. They use three layers of convolution to achieve local feature learning of residuals, simplifying the feature aggregation operation. They believed that although the feature extraction connection achieved using RFDB via 1 × 1 convolutional operations and cascade operations could effectively reduce the number of parameters, it would seriously slow down the inference speed. Therefore, they forgo multiple feature extraction connections and use several 3 × 3 conv and ReLU layers for local feature extraction. And they add the final output features to the shallow extracted features from the very first input after multiple local feature extraction. The obtained features are next passed through the 1 × 1 convolutional layer and its subsequent ESA module to obtain the final output of RLFB, as shown in Figure 14. In addition, to further reduce the runtime, the number of convolutional layers in each ConvGroups in ESA is reduced to one, which not only prevents performance degradation but also optimizes the inference time and model parameters.

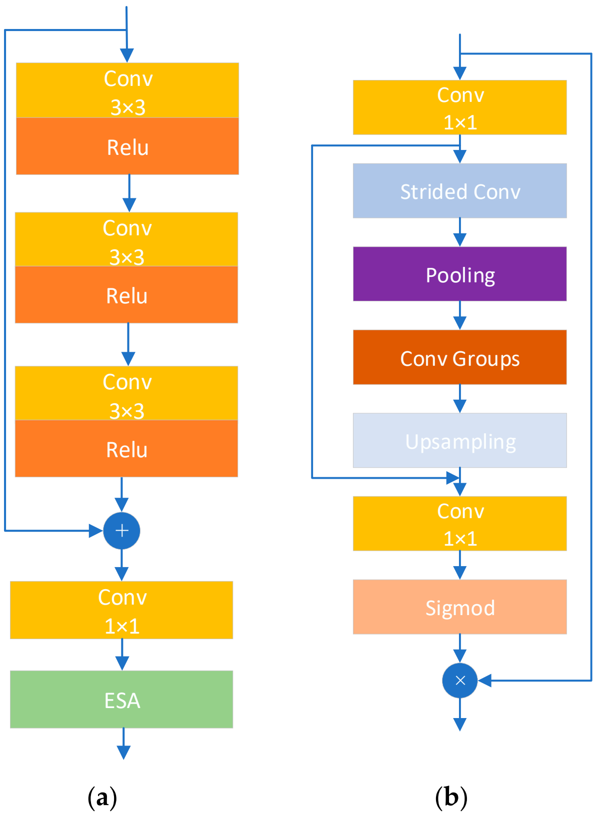


Figure 14. RLFB (**a**) and ESA (**b**) in RLFN.

5.3. Perceptual Extreme SISR

There is a balance between the two types of evaluation metrics, one focusing on picture quality and the other on perceptual quality, and there is no method yet to achieve the best picture quality while at the same time achieving optimal perceptual quality. Existing SISR methods tend to focus on the image quality of the reconstructed images; however, the perceived quality of the images is also an important indicator of the merit of the reconstructed images. In addition, the problem of SISR regarding the magnification factor at large scales has received little attention. Therefore, it is a worthwhile problem to realize large-scale SISR and to reconstruct images with excellent perceptual quality. The perceptual extreme super-resolution challenge aims to achieve super-resolution reconstruction with very large magnification factors, similar to $\times 16$ [38,68,73], as well as to achieve super-resolution reconstruction of images with high perceptual quality. In Table 8, we show the perceptual extreme SISR challenge track and winner for the period of 2017–2022, and in Table 9, we show the results of perceptual extreme SISR challenge winners for 2017–2022.

Table 8. Perceptual extreme SISR challenge track in 2017–2022.

Challenge	Track	Winner
PIRM 2018	The task divides the perceptual distortion plane into three regions in terms of RMSE, and the participant's goal is to obtain the best average perceptual quality on each perceptual plane region. Region 1 (low RMSE, high PSNR) Region 2 (middle RMSE, middle PSNR) Region 3 (high RMSE, low PSNR)	EPSR DBPN ESRGAN
AIM 2019	Track 1: The aim is to generate high-fidelity results with an amplification factor of $\times 16$. Track 2: Designed to generate high perceptual quality results with a magnification factor $\times$ 16.	DSSR MGBPv2
NTIRE 2020	Realization of amplification factor $\times$ 16	RFB-SRGAN

Table 9. Results of the perceptual extreme SISR challenge winners.

SISR Networks	PSNR	SSIM	LPIPS	PI	RMSE	TIME
EPSR	-	-	-	2.709	11.48	-
DBPN	-	-	-	2.199	12.40	-
ESRGAN	-	-	-	1.978	15.30	-
DSSR	26.79	0.7289	-	-	-	30
MGBPv2	25.44	0.6551	-	-	-	47.11
RFB-SRGAN	23.38	0.5504	0.348	3.977	-	8.1

5.3.1. EPSR/DBPN/ESRGAN (Winner of PIRM2018)

The 2018 PIRM challenge was to achieve factor $\times$ 4, the super-resolution of a single image with bicubic downsampling. Unlike in the past, this challenge aims to reconstruct perceptually good quality images. The task divides the perceptual distortion plane into three regions in terms of RMSE, and the participant's goal is to obtain the best average perceptual quality on each perceptual plane region. The basic architecture of the networks used by the winners of this challenge is all GAN networks, where the ESRGAN [69] proposed by Wang et al. achieves the best average perceptual quality over region three, as shown in Figure 15, for the specific network structure. The EPSR [74] proposed by Vasu, S. et al. obtained the best average perceptual quality over Region I. The specific network structure is shown in Figure 16. The network is trained using a combination of mean squared error loss, perceptual loss, and adversarial loss using EDSR as the generator.

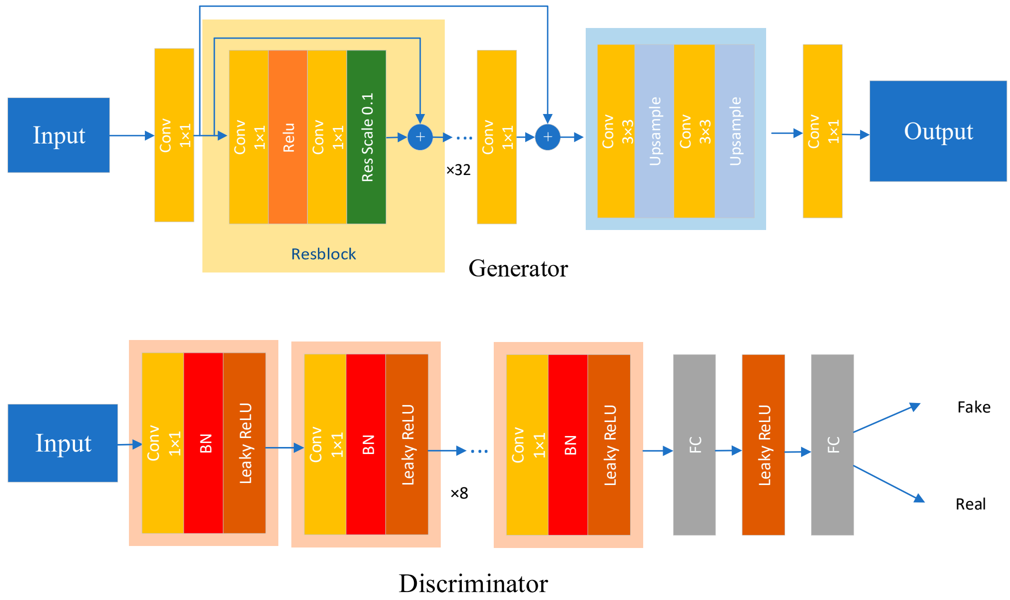


Figure 15. Network structure model of EPSR.

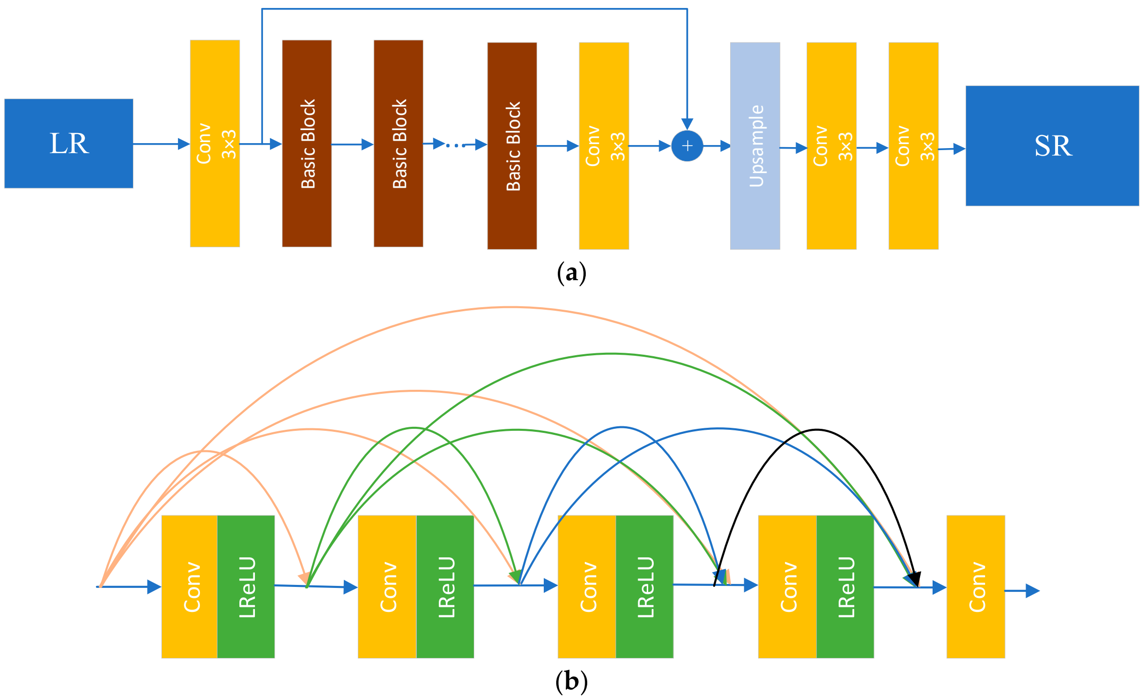


Figure 16. The overall network architecture of ESRGAN (**a**) and the basic block in it (**b**).

ESRGAN removes all BN layers compared to SRGAN [61] to train deeper networks and replaces the basic blocks in SRRESNet with RRDB to make the network easier to train. Basic blocks consist of residual modules and tight junctions, allowing more layers in the network and improving performance effectively. Meanwhile, ESRGAN uses a relative discriminator, which is no longer the probability of true and false in SRGAN, but the probability of judging the true image to be truer than the false one. And this design helps guide the generator to generate reconstructed images with more realistic texture details.

5.3.2. DSSR/MGBPv2 (Winner of AIM 2019)

AIM 2019 also presents an extreme super-resolution challenge, using the DIV8K dataset to achieve a factor $\times$ 16 super-resolution. The competition has two tracks: the first aims to generate high-fidelity results, and the second aims to generate high-perceptual-quality results.

The DSSR [75] proposed by NUAA-404 achieves superiority in generating a high-fidelity track, and the specific network structure is shown in Figure 17. The network connects two $\times$ 4 networks to achieve the target result of $\times$16. DSSR consists of two parts, i.e., SKIP and BODY. SKIP is a sub-pixel convolution module that uses the low-frequency information in LR images to reconstruct HR images. BODY consists of two networks with an amplification factor of $\times$4. The first network consists of a feature extraction layer, an ADRU layer, a GFF layer, and an AFSL layer. The second network consists of a feature extraction layer, an ADRB layer, and an AFSL layer. The BODY layer is used to reconstruct the HR image using the high-frequency information from LR, and finally, the HR image is obtained by combining the results of BODY with the results of SKIP. The ADRU module consists of four ADRB modules that are tightly connected, and the obtained features are merged via GFF. The first segment of the network consists of four ADRU modules tightly connected with the features obtained via fusion with the LFF layer. The convolution unit in ADRB consists of two wide convolutions and a Leaky ReLU, similar to WDSR, as shown in Figure 18.

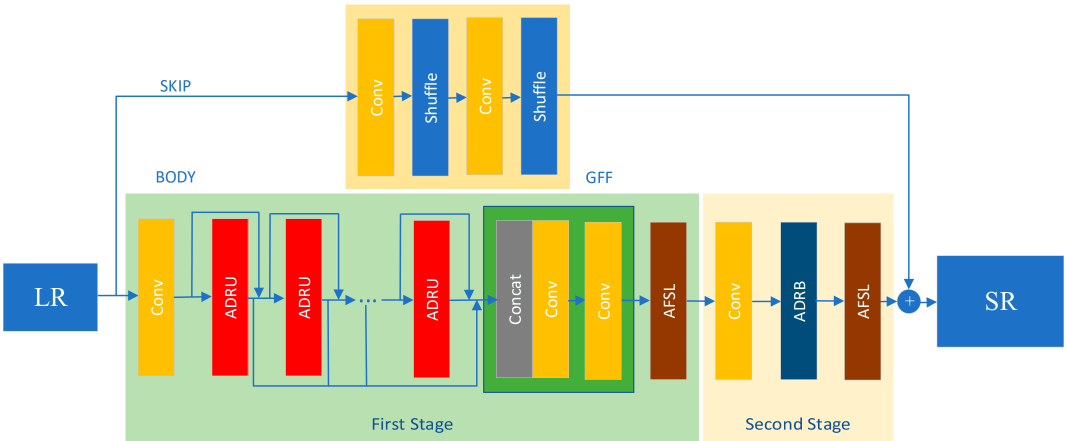


Figure 17. The overall network architecture of DSSR.

The network proposes a new reconstruction model, AFSL, which uses more parameters and more computation than the commonly used subpixel convolution, and also brings better results.

The MGBPv2 proposed by BOE-IOT-AIBD achieves superiority in generating a high sensory quality track with the specific network structure. This method combines MutiGrid (MG) and BackProjections (BP) to provide feasibility for extreme SISR tasks. Although MGBP has good results, MGBP does not work on super-resolution issues. This is mainly due to the poor quality of the reconstructed images due to the small number of parameters and the recursive network structure that causes the number of network features to remain constant along the scale. Compared to the MGBP they proposed in 2018, they made the following improvements.

The MGBPv2 uses recursive networks at the beginning of the network. And then, BOE-IOT-AIBD proposed a strategy to merge patches in inference as a way to handle large-scale images. And they simplify the main module by allowing each instance in the network to use different parameters. In addition, the team proposes a multiscale training strategy that combines distortion or perceptual loss of the output image with a reduced scale output image.

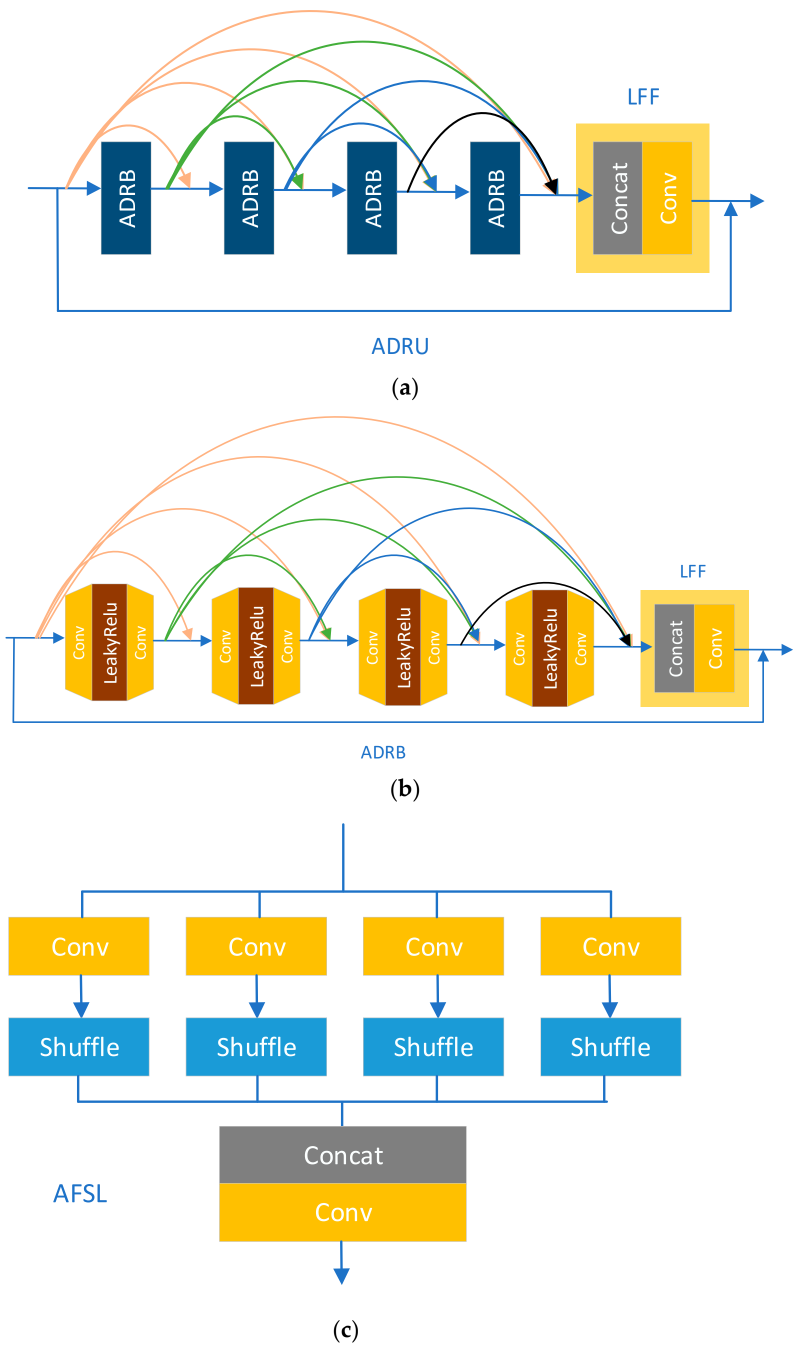


Figure 18. ADRU in DSSR (**a**), AFSL (**c**) and ADRB in ADRU (**b**).

5.3.3. RFB-SRGAN (Winner of NTIRE 2020)

The extreme super-resolution challenge is presented in NTIRE 2020, using a dataset of DIV8K, which aims to achieve a super-resolution with a magnification factor of $\times 16$. The winning model for that year is the RFB-SRGAN [76] proposed by OPPO-Research based on ESRGAN, and the specific network structure is shown in Figure 19. The network consists of five modules: shallow feature extraction module, Trunk-A, Trunk-RFB, upsampling module, and reconstruction module. Among them, the Trunk-A module consists of 16 RRDBs, and Trunk-RFB consists of 8 RFB-RDBs.

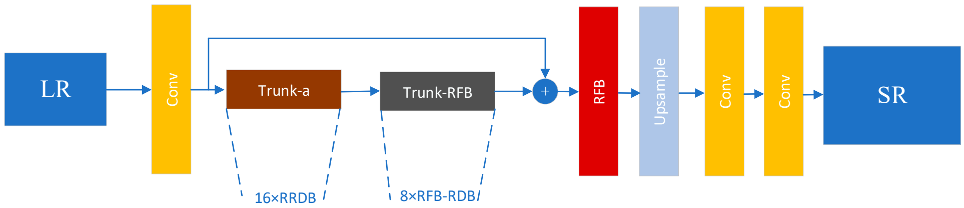
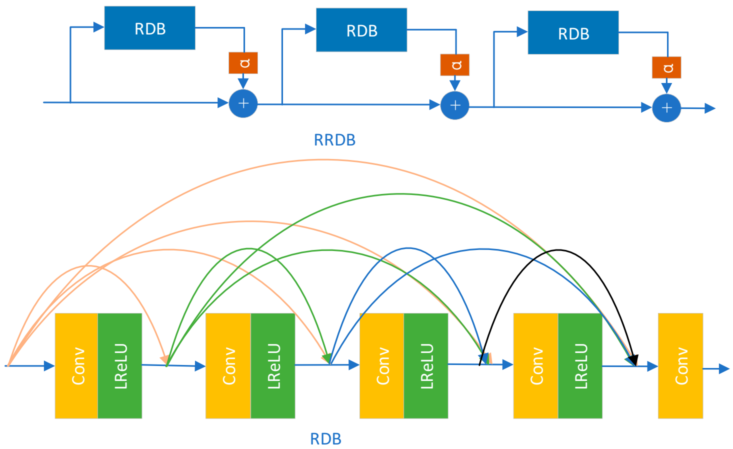
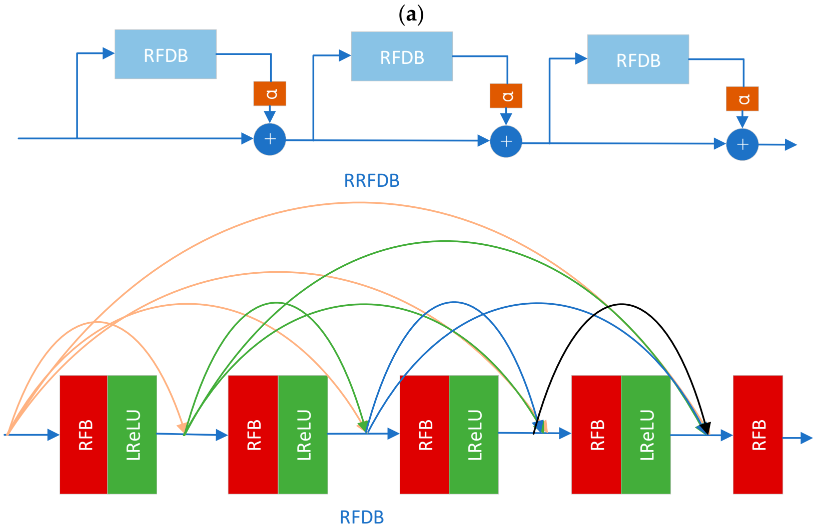


Figure 19. The overall network architecture of RFB-SRGAN.

For the perceptually extreme super-resolution challenge task, multi-scale features are needed to reconstruct the details. So, the team introduced the RFB module. The RFB module can utilize a multi-branch pool with different kernels corresponding to different sizes of receptive fields, apply its extended convolutional layers to control their eccentricity, and finally reconstruct to generate the final result. The RFB module in RFB-SRGAN uses a combination of 1×1, 1×3, and 3×1 convolutional kernels instead of large convolutional kernels such as 3×3, 5×5, and so on, as shown in Figure 20. This method effectively reduces the time and parameters needed for the computation, in addition to better extracting detailed features. The important reason for the team to use RFB is the ability to extract very detailed features that can effectively reconstruct the image.

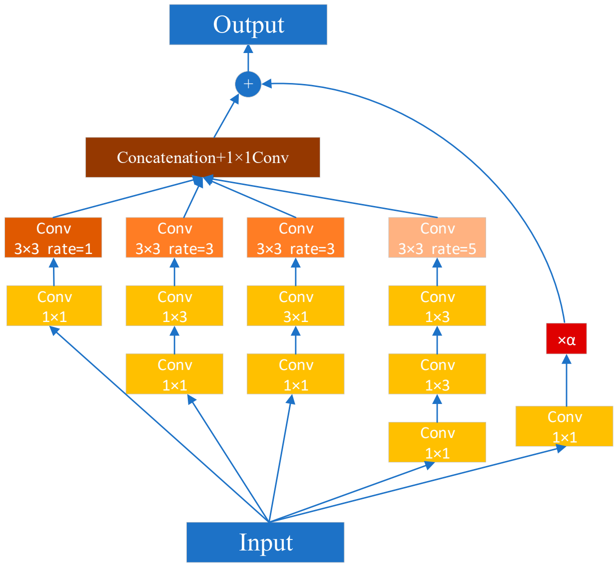


(**a**)



(**b**)

Figure 20. *Cont.*



(c)

Figure 20. RRDB (**a**), RFB-RDB (**b**), and RFB (**c**) in RFB-SRGAN.

The team also made adjustments in the upsampling section, using not only nearest interpolation (NNI) [77] or subpixel convolution (SPC) but alternating them. RFB after NNI can make the NNI transform from space to depth fully affects the depth, and RFB after SPC can make the SPC transform from depth to space fully affects space, alternating them to effectively make information exchange between space and depth. In addition to this, the use of SPC can effectively reduce the number of parameters and the running time. In Figure 21, we show upsampling method used in RFB-SRGAN.

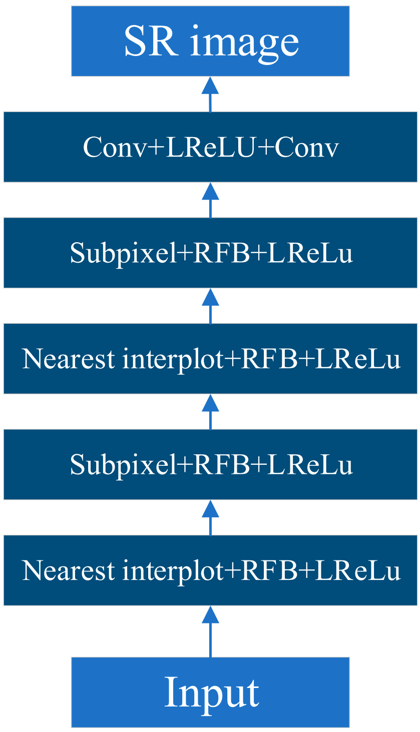


Figure 21. Upsampling method used in RFB-SRGAN.

5.4. Real-World SISR

Previous SISR network training relies on pairs of low-resolution images and high-resolution images, and the trained networks often have difficulty performing well in the real world. Since real-world LR images are degraded differently from LR images in datasets, resulting in existing SISR methods that often perform poorly on real-world images, the Real-World Image SISR Challenge aims to advance SISR models that can be used in the real world [78–80]. The challenge aims to train a network model to achieve super-resolution of natural images without paired high- and low-resolution images. In Table 10, we show the real-world SISR challenge track and winner for the period of 2017–2022, and in Table 11, we show the results of real-world SISR challenge winners for 2017–2022.

Table 10. Real-world SISR challenge track in 2017–2022.

Challenge	Track	Winner
NTIRE 2019	Track: Realization of real-world SISR.	UDSR
AIM 2019	Track 1: Designed to generate SR images with LR features, magnification factor ×4. Track 2: Designed to generate clean, high-quality HR images with a magnification factor of ×4.	DSGAN
AIM 2020	Track: The aim is to obtain images with high quality and high fidelity with magnification factors ×2, ×3, ×4.	Baidu NAS
NTIRE2020	Track 1: An unknown degradation factor is used to obtain an approximation to the real-world LR, from which the SR network is trained. Track 2: The images taken with the iPhone 3 in the OPED dataset were used as LR, from which the SR network was trained.	Real-SR

Table 11. Results of the real-world SISR challenge winners 2017–2022.

SISR Networks	PSNR	SSIM	LPIPS	MOS
UDSR	29.00	0.84	-	-
DSGAN [Track 1]	22.65	0.48	0.36	2.22
DSGAN [Track 2]	20.72	0.52	0.40	2.34
Baidu NAS [×2]	33.460	0.9237	-	-
Baidu NAS [×3]	30.950	0.876	-	-
Baidu NAS [×4]	31.396	0.875	-	-
Real SR	24.67	0.683	0.232	2.195

5.4.1. UDSR (Winner of NTIRE 2019)

The task of this challenge was to reconstruct real-world images, using the dataset RealSR, and the winner was UDSR, proposed by the SuperRior team.

In UDSR, the depth feature map is obtained from the input image via the convolution layer. The low-resolution image feature map is obtained via the residual block, and the low-resolution image feature map is used as the input. The first path processes the feature map via the residual block, the second path downsamples the feature map after going via the residual block, and the third path downsamples the feature map again. And then, the fourth path upsamples the obtained feature map and applies the residual block and convolution block. In addition, they output the highest-resolution feature maps as residual images and add them to the input images The high-resolution images obtained from the three paths are combined with the input image to achieve the final output image. We show the Network architecture of UDSR in Figure 22.

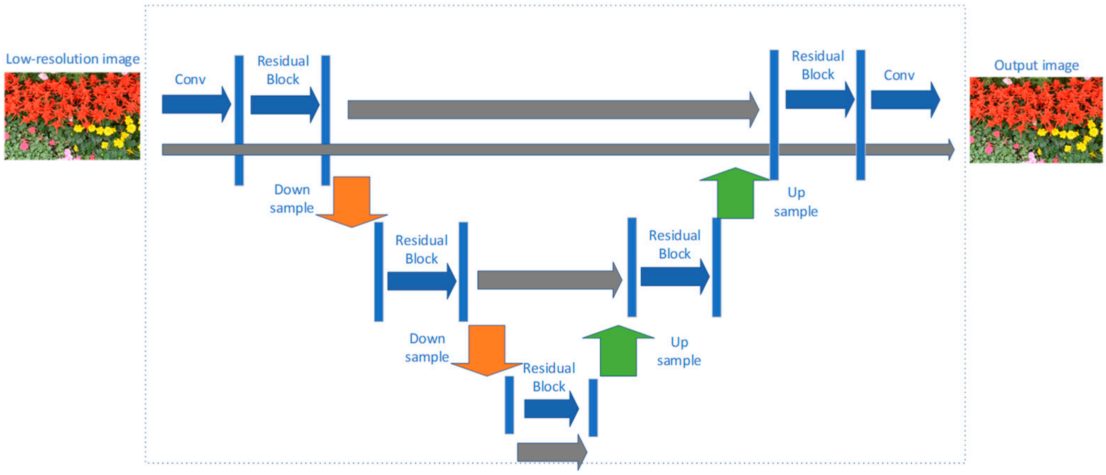


Figure 22. Network architecture of UDSR.

A tandem structure is used for training UDSR, where the HR image is first restored to its original size after quadruple downsampling the HR image of the first segment to calculate the loss function. The output of the first segment is used as the input of the second segment, and the HR image is restored to its original size after quadruple downsampling the HR image of the second segment to calculate the loss function. And the output of the second segment is used as the input of the third segment to calculate the loss function using the original HR image. After this three-stage training, the network can recover the LR image to the HR image. We show the training method of UDSR in Figure 23.

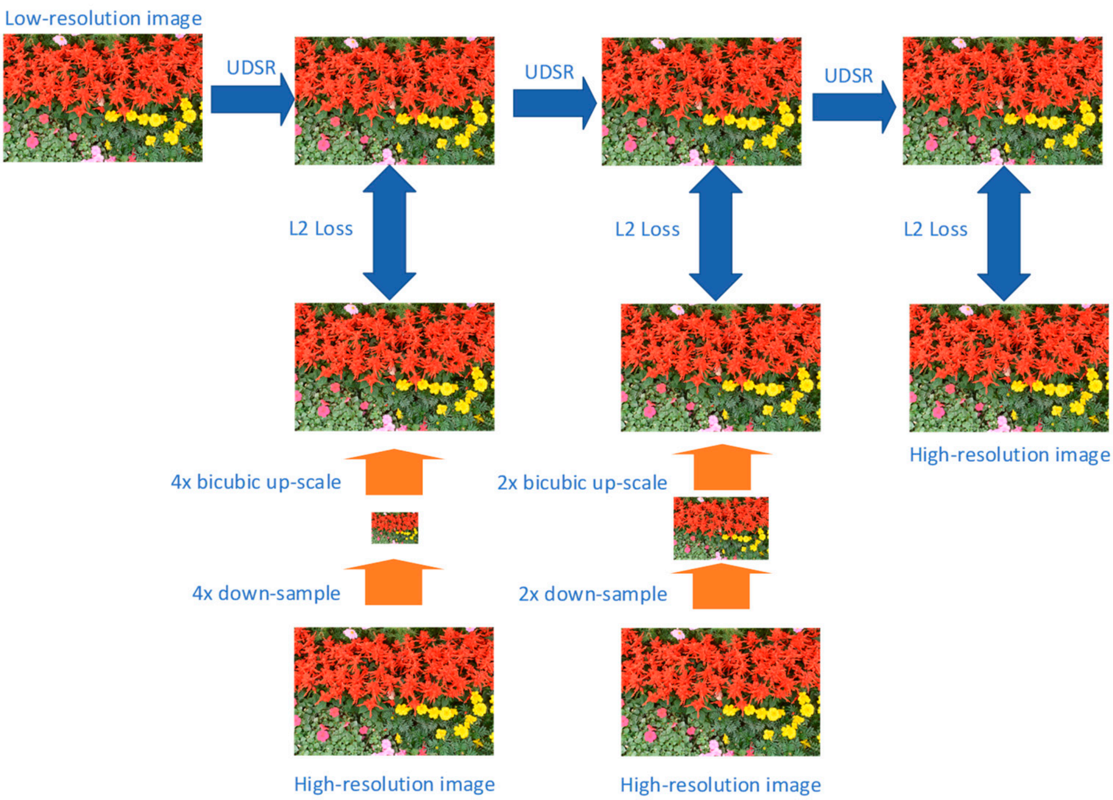


Figure 23. Training method of UDSR.

5.4.2. DSGAN (Winner of AIM 2019)

AIM 2019 presents the challenge of super-resolution in the real world with two tracks. One is to reconstruct high-resolution images with guaranteed low-quality picture features. And the other is to provide a set of unrelated images of the same quality as the target, with the learning goal of generating clean, high-quality HR images. The magnification factors of the target images are all $\times 4$.

The DSGAN [81] proposed by the Mad Demon team achieves superiority in both tracks. In Figure 24, we show the network structure. The network is divided into two phases. The first phase is to generate LR images with real-world LR image features. The second phase is to train the network in the form of supervised LR-HR pairs formed in the first phase.

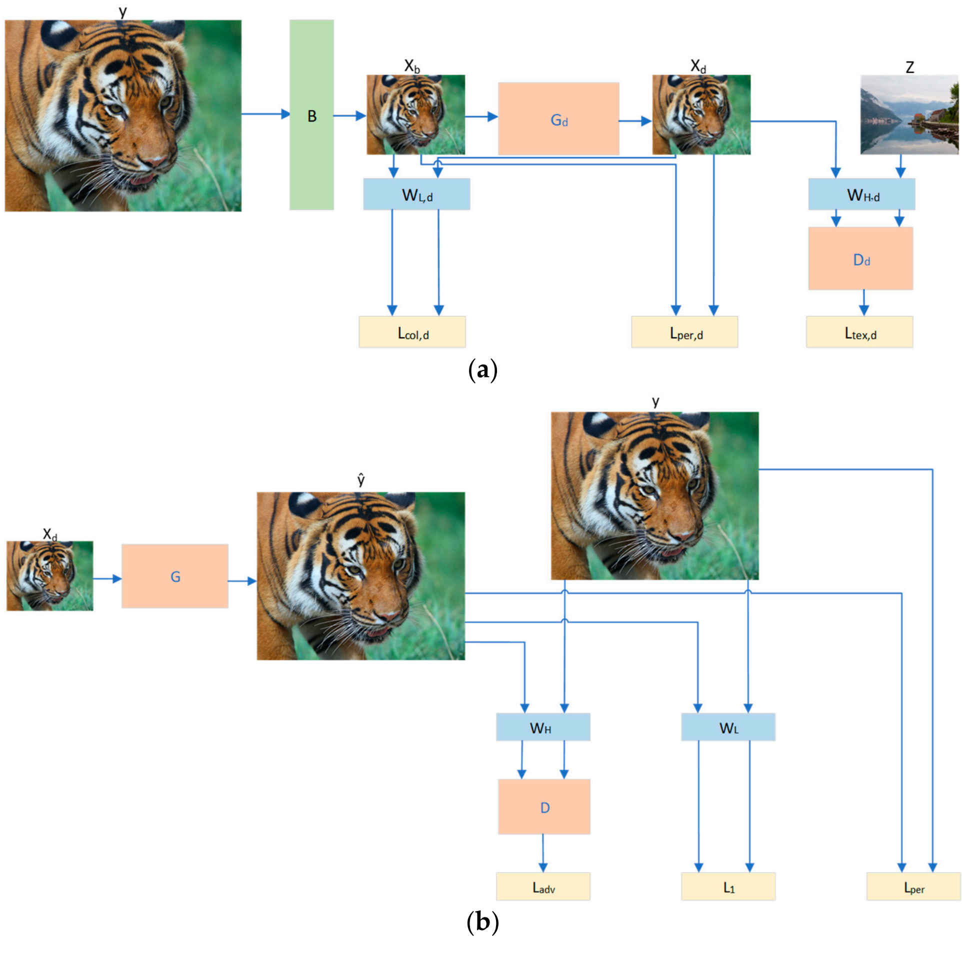
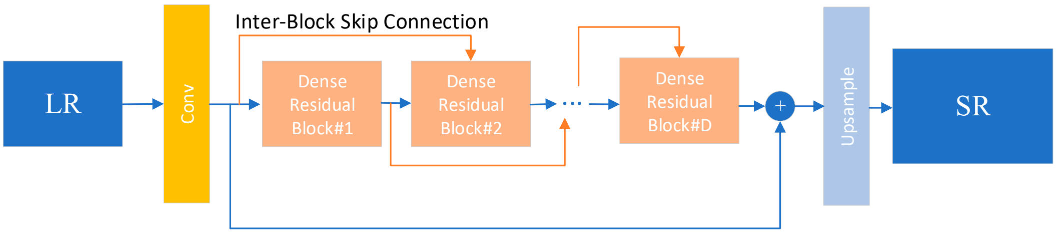


Figure 24. DSGAN generates LR images in the first stage (**a**) and pairs of LR-HR images in the second stage for training (**b**).

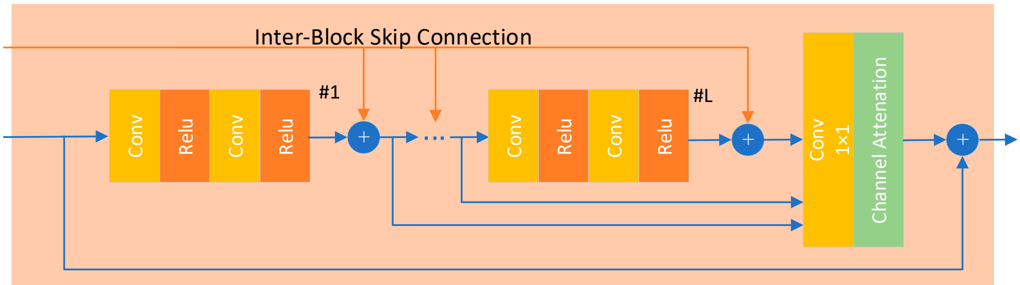
In the first stage, the HR image y is bicubic downsampled to obtain x_b, and x_b is used as the input to obtain the LR image x_d via the generator. The discriminator is used to determine which of x_d and z is the synthetic LR image and which is the real LR image. In the second stage, the SR network is trained based on the obtained image pairs. During the training, the generated SR images are separated from the low and high frequencies using filters. The low frequencies use L_1 loss to focus on the recovery of image content, and the high frequencies use the adversarial loss to focus on the recovery of image details. In addition to obtaining better perceptual quality and also better combining low- and high-frequency information, a perceptual loss function is introduced; the perceptual loss function also allows better recovery of the image content. The SR network used is ESRGAN.

5.4.3. Baidu NAS (Winner of AIM 2020)

The 2020 AIM presented a real-world super-resolution challenge using the DRealSR dataset, and the winner that year was Baidu's proposed GP-NAS-based design for super-resolution in search space [82]. Baidu focused more on the macroscopic network structure, using the GP-NAS method to search for parameters of key network structures and generate multiple alternative models. In Figure 25, we show overall architecture of the GP-NAS-based network model.



(**a**)



(**b**)

Figure 25. Overall architecture of the GP-NAS-based network model proposed by BaiDu (**a**), where the DRB (**b**). "#1" means the first double-layer convolutions. "#L" means the Lth double-layer convolutions.

The backbone model of this method is DRBN, and the whole network consists of DRB except for the shallow feature extraction convolution and the end upsampling module. The shallow feature extraction convolution converts the input three-channel image into F-channel shallow features. Each DRB consists of L double-layer convolutions, and the L outputs in the DRB are connected at the end with a 1×1 convolution and passed through a channel attention module. There are two types of jump connections in each DRB, intra-block jump connections and extra-block jump connections. There are three key hyperparameters in this network: F is the number of channels, D is the number of DRB layers in the network, and L is the number of bilayer convolutions in each DRB.

While previous works have often used expertise or experience to make choices about these hyperparameters, the Baidu team used a neural architecture search based on a Gaussian process to determine these hyperparameters as a way to obtain a network architecture with optimal performance. The method combines AI and super-resolution networks, offering new possibilities for the development of super-resolution networks.

5.4.4. Real SR (Winner of NTIRE 2020)

NTIRE 2020 proposes the real-world resolution challenge, which is divided into two tracks, one using an unknown degradation factor to obtain LR and the other using iPhone 3 images from the OPED dataset, both aiming to obtain the best perceptual quality images.

The model proposed by Impressionism achieved superiority in both tracks [83]. In Figure 26, we show the network used by RealSR. The team designs a new degradation model for real-world images by estimating degradation kernels and blur kernels and proposes a new real-world super-resolution model with the aim of better perceptual quality.

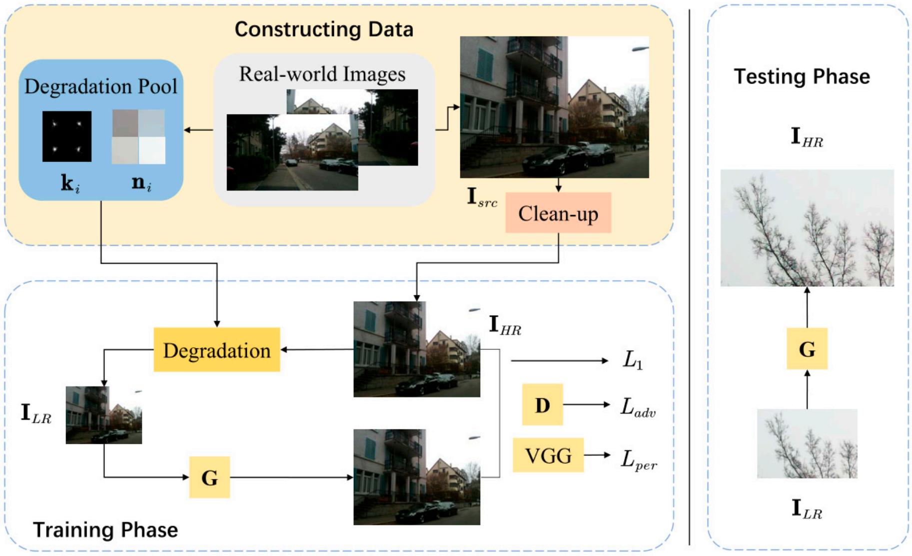


Figure 26. Network model used by Real SR.

The degradation model uses kernel estimation similar to Kernel GAN [84] to estimate degradation kernels from real-world images. To make the degraded image and the source image have the same noise distribution, the team obtains noise directly from the real-world image, and the team constructs a degradation pool from the degradation kernels and the obtained noise. To obtain more HR images, the team bicubic downsamples the real-world images to remove the noise to obtain clean images, degrades the obtained clean images with randomly selected blur kernels from the degradation pool to obtain LR images with noise and blur similar to the real-world ones, and finally trains the SR network.

The team designed the SR model based on ESRGAN. The generator adopts the structure of RRDB, and the loss function uses the weighted sum of L1, perceptual loss, and adversarial loss. The perceptual loss uses the inactive features of VGG-19 to enhance the low-frequency features; the adversarial loss is used to enhance the details of the image to make the image look more realistic. In addition, the discriminator uses a patch discriminator instead of VGG-128 for two reasons.

Firstly, VGG-128 can only discriminate images of size 128, which does not perform well on multi-scale tasks. Secondly, VGG-128 has a deeper network, which focuses more on global features and ignores local features, while the patch discriminator has a fixed perceptual field due to its fully convolutional structure, and its output values are only related to the local part, and the local loss is fed back to the generator to optimize the local details. To ensure overall consistency, the final error uses the average of all local errors.

6. Conclusions

In this paper, we present an overview of the challenge tasks on SISR for the period of 2017–2022. In Section 2, we discuss the datasets used in previous years' challenges, using different datasets to meet the requirements of each challenge task and to provide enough a priori information to improve the efficiency of the network. Section 3 introduces the IQA methods commonly used in previous years' competitions. Section 4 shows the challenge tasks and the network architectures of the winners in previous years. The challenge tasks in previous years can be broadly classified into four categories: 1. Classic SISR challenge, including two tracks of known bicubic downsampling degradation and unknown degradation factor super-resolution; 2. Perceptual Extreme SISR challenge, mainly to achieve super-resolution with special amplification factor $\times$ 16; 3. Efficient SISR challenge, aiming to achieve reduced network inference time and a number of computations; 4. Real-world SISR challenge, which aims to advance the development of networks with super-resolution that also work in the real world. Among these challenges, many effective approaches have been proposed and applied to their network architectures to improve network performance, such as attention modules, back-projection networks, tight junctions, residual networks, information distillation, recursive networks, and the recently acclaimed Transformer. Despite the advancements in SISR with the involvement of deep-learning-based methods, there are still a number of challenges that need to be considered. We present an outlook of future work in the following items.

1. Normalization layer: During 2017–2022, many superior networks used different normalization layers to improve network performance, such as EDSR, to remove the BN layer, TCIR to use the LR layer, etc. The BN layer normalizes the same batch of data, which can accelerate network convergence, control overfitting, allow the use of larger learning rates, and is more applicable to scenarios with larger batch sizes. The LN layer normalizes the data of the whole layer and is insensitive to the size of the batch, in addition to inheriting the advantages of the BN layer. Therefore, it is often necessary to select the appropriate normalization layer by experience when designing the network. Switchable Normalization (SN) [85] was proposed in 2018, combining various operations of IN, LN, and BN to select the appropriate normalization layer for the network when targeting different vision tasks. This may become one of the normalization methods often used in SISR tasks in the future.
2. More efficient or lighter networks: Using CNN networks to implement SR is fast and occupies less memory, but some edge information will be lost; using Transformer networks to implement SR can be achieved using full-text information to reconstruct images, but it is slower and occupies more memory. In addition, although CNN networks have advantages in local feature extraction, they are still inadequate for global feature representation. Transformer, on the other hand, has a good sense of global features but ignores local feature details. In recent years, many networks combining CNN and Transformer have been proposed. TCIR is a typical network combining CNN and Transformer. This network added several RRDB modules to the basic module of TCIR, which combined the advantages of both CNN and Transformer and achieved first place in the AIM 2022 compressed image super-resolution challenge. So, further research can be conducted in this direction to design networks to better combine the advantages of both.
3. The need for more accurate and effective IQA: Existing IQA methods are difficult to balance perceptual quality and image quality, and images that score high in image quality often do not score high in perceptual quality. Therefore, we need a more suitable IQA method to evaluate both perceptual quality and image quality.
4. Unsupervised real-world image super-resolution network model training: The currently proposed super-resolution challenge superiority methods on real-world images are based on learning degradation methods in the real world, from which LR images corresponding to HR are obtained, and then pairwise supervised network training is performed to obtain the network model. The performance of the obtained SISR network depends more on the ability to generate LR images with similar blurring as real-world LR images for training. Such networks are also often not strongly generalizable due to the various reasons for the blurring of real-world images. Therefore, how to implement unsupervised super-resolution training on real-world images is a direction for future development.

Author Contributions: Conceptualization, S.Y. and C.X.; methodology, S.Y. and C.X.; software, S.Y.; validation, S.Y., S.Z. and Y.H.; formal analysis, S.Z.; investigation, Y.H.; resources, C.X.; data curation, S.Y.; writing—original draft preparation, S.Y.; writing—review and editing, S.Y. and C.X.; visualization, S.Y. and C.X.; supervision, C.X.; project administration, C.X.; funding acquisition, C.X. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded in part by the National Natural Science Foundation of China under Grant 61901221 and 62203012, in part by the Postgraduate Research and Practice Innovation Program of Jiangsu Province under Grant KYCX21_0872, and in part by the National Key Research and Development Program of China under Grant 2019YFD1100404.

Data Availability Statement: The datasets analyzed during this current study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, Q.; Song, H.; Yu, J.; Kim, K. Current Development and Applications of Super-Resolution Ultrasound Imaging. *Sensors* **2021**, *21*, 2417. [CrossRef]
2. Lin, F.; Rojas, J.D.; Dayton, P.A. Super resolution contrast ultrasound imaging: Analysis of imaging resolution and application to imaging tumor angiogenesis. In Proceedings of the 2016 IEEE International Ultrasonics Symposium (IUS), Tours, France, 18–21 September 2016; pp. 1–4. [CrossRef]
3. Mahapatra, D.; Bozorgtabar, B.; Garnavi, R. Image super-resolution using progressive generative adversarial networks for medical image analysis. *Comput. Med. Imaging Graph.* **2018**, *71*, 30–39. [CrossRef]
4. Xie, C.; Zhu, H.; Fei, Y. Deep coordinate attention network for single image super-resolution. *IET Image Process.* **2021**, *16*, 273–284. [CrossRef]
5. Zhang, L.; Zhang, H.; Shen, H.; Li, P. A super-resolution reconstruction algorithm for surveillance images. *Signal Process.* **2010**, *90*, 848–859. [CrossRef]
6. Rasti, P.; Uiboupin, T.; Escalera, S.; Anbarjafari, G. Convolutional Neural Network Super Resolution for Face Recognition in Surveillance Monitoring. In Proceedings of the Articulated Motion and Deformable Objects: 9th International Conference, AMDO 2016, Palma de Mallorca, Spain, 13–15 July 2016; pp. 175–184.
7. Mudunuri, S.P.; Biswas, S. Low Resolution Face Recognition across Variations in Pose and Illumination. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *38*, 1034–1040. [CrossRef]
8. Wang, P.; Bayram, B.; Sertel, E. A comprehensive review on deep learning based remote sensing image super-resolution methods. *Earth-Sci. Rev.* **2022**, *232*, 104110. [CrossRef]
9. Lillesand, T.; Kiefer, R.W.; Chipman, J. *Remote Sensing and Image Interpretation*, 5th ed.; John Wiley & Sons: Hobokan, NJ, USA, 2004; ISBN 0471152277.
10. Lei, S.; Shi, Z.; Zou, Z. Super-Resolution for Remote Sensing Images via Local–Global Combined Network. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1243–1247. [CrossRef]
11. Huang, Y.; Shao, L.; Frangi, A.F. Simultaneous Super-Resolution and Cross-Modality Synthesis of 3D Medical Images Using Weakly-Supervised Joint Convolutional Sparse Coding. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 5787–5796. [CrossRef]
12. Isaac, J.S.; Kulkarni, R. Super resolution techniques for medical image processing. In Proceedings of the 2015 International Conference on Technologies for Sustainable Development (ICTSD), Mumbai, India, 4–6 February 2015; pp. 1–6.
13. Robinson, M.D.; Chiu, S.J.; Toth, C.A.; Izatt, J.A.; Lo, J.Y.; Farsiu, S. *New Applications of Super-Resolution in Medical Imaging*; CRC Press: Boca Raton, FL, USA, 2017; pp. 383–412. [CrossRef]
14. Dai, D.; Wang, Y.; Chen, Y.; Van Gool, L. Is image super-resolution helpful for other vision tasks? In Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Placid, NY, USA, 7–10 March 2016; pp. 1–9.
15. Haris, M.; Shakhnarovich, G.; Ukita, N. Task-Driven Super Resolution: Object Detection in Low-Resolution Images. In Proceedings of the Neural Information Processing: 28th International Conference, ICONIP 2021, Bali, Indonesia, 8–12 December 2021; pp. 387–395. [CrossRef]
16. Sajjadi, M.S.M.; Scholkopf, B.; Hirsch, M. Enhancenet: Single image super-resolution through automated texture synthesis. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 4491–4500.
17. Bai, Y.; Zhang, Y.; Ding, M.; Ghanem, B. SOD-MTGAN: Small Object Detection via Multi-Task Generative Adversarial Network. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 206–221. [CrossRef]
18. Guo, Z.; Wu, G.; Song, X.; Yuan, W.; Chen, Q.; Zhang, H.; Shi, X.; Xu, M.; Xu, Y.; Shibasaki, R.; et al. Super-Resolution Integrated Building Semantic Segmentation for Multi-Source Remote Sensing Imagery. *IEEE Access* **2019**, *7*, 99381–99397. [CrossRef]
19. Zhao, C.; Shao, M.; Carass, A.; Li, H.; Dewey, B.E.; Ellingsen, L.M.; Woo, J.; Guttman, M.A.; Blitz, A.M.; Stone, M.; et al. Applications of a deep learning method for anti-aliasing and super-resolution in MRI. *Magn. Reson. Imaging* **2019**, *64*, 132–141. [CrossRef]
20. Xie, C.; Zeng, W.; Lu, X. Fast Single-Image Super-Resolution via Deep Network With Component Learning. *IEEE Trans. Circuits Syst. Video Technol.* **2018**, *29*, 3473–3486. [CrossRef]
21. Wang, L.; Li, D.; Zhu, Y.; Tian, L.; Shan, Y. Dual super-resolution learning for semantic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 3774–3783.

22. Pang, Y.; Cao, J.; Wang, J.; Han, J. JCS-Net: Joint Classification and Super-Resolution Network for Small-Scale Pedestrian Detection in Surveillance Images. *IEEE Trans. Inf. Forensics Secur.* **2019**, *14*, 3322–3331. [CrossRef]
23. Wang, Z.; Chang, S.; Yang, Y.; Liu, D.; Huang, T.S. Studying very low resolution recognition using deep networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 4792–4800.
24. Yang, X.; Wu, W.; Liu, K.; Kim, P.W.; Sangaiah, A.K.; Jeon, G. Long-Distance Object Recognition with Image Super Resolution: A Comparative Study. *IEEE Access* **2018**, *6*, 13429–13438. [CrossRef]
25. Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.; Villena-Martinez, V.; Martinez-Gonzalez, P.; Garcia-Rodriguez, J. A survey on deep learning techniques for image and video semantic segmentation. *Appl. Soft Comput.* **2018**, *70*, 41–65. [CrossRef]
26. Bouwmans, T.; Javed, S.; Sultana, M.; Jung, S.K. Deep neural network concepts for background subtraction: A systematic review and comparative evaluation. *Neural Netw.* **2019**, *117*, 8–66. [CrossRef] [PubMed]
27. Yan, X.; Liu, Y.; Huang, D.; Jia, M. A new approach to health condition identification of rolling bearing using hierarchical dispersion entropy and improved Laplacian score. *Struct. Health Monit.* **2020**, *20*, 1169–1195. [CrossRef]
28. Jiang, D.; Wang, M.; Sun, Y.; Hang, X. Equivalent Modeling of Bolted Connections under Transverse Load Using Iwan-Based Material Properties. *Metals* **2023**, *13*, 91. [CrossRef]
29. Wang, Y.; Huang, Z.; Zhu, P.; Zhu, R.; Hu, T.; Zhang, D.; Jiang, D. Effects of compressed speckle image on digital image correlation for vibration measurement. *Measurement* **2023**, *217*, 113041. [CrossRef]
30. Jiang, D.; Wang, Y.; Hu, J.; Qian, H.; Zhu, R. Automatic modal identification based on similarity filtering and fuzzy clustering. *J. Vib. Control* **2023**. [CrossRef]
31. Liu, A.; Liu, Y.; Gu, J.; Qiao, Y.; Dong, C. Blind image super-resolution: A survey and beyond. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 5461–5480. [CrossRef]
32. Wang, Z.; Chen, J.; Hoi, S.C. Deep learning for image super-resolution: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 3365–3387. [CrossRef]
33. Yang, W.; Zhang, X.; Tian, Y.; Wang, W.; Xue, J.-H.; Liao, Q. Deep Learning for Single Image Super-Resolution: A Brief Review. *IEEE Trans. Multimed.* **2019**, *21*, 3106–3121. [CrossRef]
34. Li, J.; Pei, Z.; Zeng, T. From beginner to master: A survey for deep learning-based single-image super-resolution. *arXiv* **2021**, arXiv:2109.14335.
35. Lepcha, D.C.; Goyal, B.; Dogra, A.; Goyal, V. Image Super-resolution: A Comprehensive Review, Recent Trends, Challenges and Applications. *Inf. Fusion* **2022**, *91*, 230–260. [CrossRef]
36. Nasrollahi, K.; Moeslund, T.B. Super-resolution: A comprehensive survey. *Mach. Vis. Appl.* **2004**, *25*, 1423–1468. [CrossRef]
37. Timofte, R.; Agustsson, E.; Van Gool, L.; Yang, M.-H.; Zhang, L. Ntire 2017 challenge on single image super-resolution: Methods and results. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 114–125.
38. Blau, Y.; Mechrez, R.; Timofte, R.; Michaeli, T.; Zelnik-Manor, L. The 2018 PIRM challenge on perceptual image super-resolution. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018.
39. Zhang, K.; Danelljan, M.; Li, Y.; Timofte, R.; Liu, J.; Tang, J.; Wu, G.; Zhu, Y.; He, X.; Xu, W.; et al. AIM 2020 Challenge on Efficient Super-Resolution: Methods and Results. In Proceedings of the Computer Vision–ECCV 2020 Workshops, Glasgow, UK, 23–28 August 2020; pp. 5–40. [CrossRef]
40. Lugmayr, A.; Danelljan, M.; Timofte, R.; Fritsche, M.; Gu, S.; Purohit, K.; Kandula, P.; Suin, M.; Rajagoapalan, A.N.; Joon, N.H.; et al. AIM 2019 Challenge on Real-World Image Super-Resolution: Methods and Results. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Seoul, Republic of Korea, 27–28 October 2019; pp. 3575–3583. [CrossRef]
41. Yang, J.; Wright, J.; Huang, T.; Ma, Y. Image super-resolution as sparse representation of raw image patches. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AL, USA, 23–28 June 2008; pp. 1–8.
42. Bevilacqua, M.; Roumy, A.; Guillemot, C.; Alberi Morel, M.-L. Low-Complexity Single-Image SUPER-RESolution based on Nonnegative Neighbor Embedding. In Proceedings of the British Machine Vision Conference, Guildford, UK, 3–7 September 2012.
43. Zeyde, R.; Elad, M.; Protter, M. On Single Image Scale-Up Using Sparse-Representations. In Proceedings of the Curves and Surfaces: 7th International Conference, Avignon, France, 24–30 June 2010; pp. 711–730. [CrossRef]
44. Timofte, R.; De Smet, V.; Van Gool, L. A+: Adjusted Anchored Neighborhood Regression for Fast Super-Resolution. In *Asian Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 111–126. [CrossRef]
45. Huang, J.-B.; Singh, A.; Ahuja, N. Single image super-resolution from transformed self-exemplars. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 5197–5206.
46. Timofte, R.; De Smet, V.; Van Gool, L. Anchored Neighborhood Regression for Fast Example-Based Super-Resolution. In Proceedings of the 2013 IEEE International Conference on Computer Vision, Sydney, Australia, 1–8 December 2013; pp. 1920–1927.
47. Agustsson, E.; Timofte, R. Ntire 2017 challenge on single image super-resolution: Dataset and study. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 126–135.
48. Cai, J.; Zeng, H.; Yong, H.; Cao, Z.; Zhang, L. Toward Real-World Single Image Super-Resolution: A New Benchmark and a New Model. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019. [CrossRef]

49. Wei, P.; Xie, Z.; Lu, H.; Zhan, Z.; Ye, Q.; Zuo, W.; Lin, L. Component Divide-and-Conquer for Real-World Image Super-Resolution. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; pp. 101–117. [CrossRef]
50. Ignatov, A.; Kobyshev, N.; Timofte, R.; Vanhoey, K. DSLR-Quality Photos on Mobile Devices with Deep Convolutional Networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 3297–3305. [CrossRef]
51. Gu, S.; Lugmayr, A.; Danelljan, M.; Fritsche, M.; Lamour, J.; Timofte, R. Div8k: Diverse 8k resolution image dataset. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Seoul, Republic of Korea, 27–28 October 2019; pp. 3512–3516.
52. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef]
53. Ma, C.; Yang, C.-Y.; Yang, X.; Yang, M.-H. Learning a no-reference quality metric for single-image super-resolution. *Comput. Vis. Image Underst.* **2017**, *158*, 1–16. [CrossRef]
54. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a "completely blind" image quality analyzer. *IEEE Signal Process. Lett.* **2012**, *20*, 209–212. [CrossRef]
55. Zhang, R.; Isola, P.; Efros, A.A.; Shechtman, E.; Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 586–595.
56. Mittal, A.; Moorthy, A.K.; Bovik, A.C. No-Reference Image Quality Assessment in the Spatial Domain. *IEEE Trans. Image Process.* **2012**, *21*, 4695–4708. [CrossRef]
57. Venkatanath, N.; Praneeth, D.; Bh, M.C.; Channappayya, S.S.; Medasani, S.S. Blind image quality evaluation using perception based features. In Proceedings of the 2015 Twenty First National Conference on Communications (NCC), Mumbai, India, 27 February–1 March 2015; pp. 1–6.
58. Timofte, R.; Gu, S.; Wu, J.; Van Gool, L. Ntire 2018 challenge on single image super-resolution: Methods and results. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 852–863.
59. Yang, R.; Timofte, R.; Li, X.; Zhang, Q.; Zhang, L.; Liu, F.; He, D.; Li, F.; Zheng, H.; Yuan, W. Aim 2022 challenge on super-resolution of compressed image and video: Dataset, methods and results. In Proceedings of the Computer Vision–ECCV 2022 Workshops, Tel Aviv, Israel, 23–27 October 2022; pp. 174–202.
60. Lim, B.; Son, S.; Kim, H.; Nah, S.; Mu Lee, K. Enhanced deep residual networks for single image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 136–144.
61. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.P.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017.
62. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 448–456.
63. Haris, M.; Shakhnarovich, G.; Ukita, N. Deep back-projection networks for super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 1664–1673.
64. Yu, J.; Fan, Y.; Yang, J.; Xu, N.; Wang, Z.; Wang, X.; Huang, T. Wide activation for efficient and accurate image super-resolution. *arXiv* **2018**, arXiv:1808.08718.
65. Salimans, T.; Kingma, D.P. Weight normalization: A simple reparameterization to accelerate training of deep neural networks. In Proceedings of the 30th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain, 5–10 December 2016.
66. Liang, J.; Cao, J.; Sun, G.; Zhang, K.; Van Gool, L.; Timofte, R. Swinir: Image restoration using swin transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 10–17 October 2021; pp. 1833–1844.
67. Li, Y.; Zhang, K.; Timofte, R.; Van Gool, L.; Kong, F.; Li, M.; Liu, S.; Du, Z.; Liu, D.; Zhou, C. NTIRE 2022 challenge on efficient super-resolution: Methods and results. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 1062–1102.
68. Gu, S.; Danelljan, M.; Timofte, R.; Haris, M.; Akita, K.; Shakhnarovic, G.; Ukita, N.; Michelini, P.N.; Chen, W.; Liu, H. Aim 2019 challenge on image extreme super-resolution: Methods and results. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Seoul, Republic of Korea, 27–28 October 2019; pp. 3556–3564.
69. Wang, X.; Yu, K.; Wu, S.; Gu, J.; Liu, Y.; Dong, C.; Qiao, Y.; Change Loy, C. Esrgan: Enhanced super-resolution generative adversarial networks. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018.
70. Hui, Z.; Gao, X.; Yang, Y.; Wang, X. Lightweight image super-resolution with information multi-distillation network. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; pp. 2024–2032.

71. Liu, J.; Tang, J.; Wu, G. Residual feature distillation network for lightweight image super-resolution. In *Computer Vision—ECCV 2020 Workshops, Proceedings of the ECCV 2020: European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020*; Springer: Cham, Switzerland, 2020; pp. 41–55.
72. Kong, F.; Li, M.; Liu, S.; Liu, D.; He, J.; Bai, Y.; Chen, F.; Fu, L. Residual local feature network for efficient super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–20 June 2022; pp. 766–776.
73. Zhang, K.; Gu, S.; Timofte, R. Ntire 2020 challenge on perceptual extreme super-resolution: Methods and results. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 492–493.
74. Vasu, S.; Madam, N.T.; Rajagopalan, A.N. Analyzing Perception-Distortion Tradeoff Using Enhanced Perceptual Super-Resolution Network. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2019; Springer: Cham, Switzerland, 2019; pp. 114–131. [CrossRef]
75. Xie, T.; Yang, X.; Jia, Y.; Zhu, C.; Xiaochuan, L. Adaptive densely connected single image super-resolution. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Seoul, Republic of Korea, 27–28 October 2019; pp. 3432–3440.
76. Shang, T.; Dai, Q.; Zhu, S.; Yang, T.; Guo, Y. Perceptual extreme super-resolution network with receptive field block. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 440–441.
77. Rukundo, O.; Cao, H. Nearest neighbor value interpolation. *arXiv* **2012**, arXiv:1211.1768.
78. Wei, P.; Lu, H.; Timofte, R.; Lin, L.; Zuo, W.; Pan, Z.; Li, B.; Xi, T.; Fan, Y.; Zhang, G.; et al. AIM 2020 Challenge on Real Image Super-Resolution: Methods and Results. In Proceedings of the Computer Vision–ECCV 2020 Workshops, Glasgow, UK, 23–28 August 2020; pp. 392–422. [CrossRef]
79. Cai, J.; Gu, S.; Timofte, R.; Zhang, L. Ntire 2019 challenge on real image super-resolution: Methods and results. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 15–20 June 2019.
80. Lugmayr, A.; Danelljan, M.; Timofte, R. Ntire 2020 challenge on real-world image super-resolution: Methods and results. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 494–495.
81. Wang, W.; Zhang, H.; Yuan, Z.; Wang, C. Unsupervised Real-World Super-Resolution: A Domain Adaptation Perspective. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 4298–4307. [CrossRef]
82. Pan, Z.; Li, B.; Xi, T.; Fan, Y.; Zhang, G.; Liu, J.; Han, J.; Ding, E. Real Image Super Resolution via Heterogeneous Model Ensemble Using GP-NAS. In Proceedings of the Computer Vision–ECCV 2020 Workshops, Glasgow, UK, 23–28 August 2020; pp. 423–436. [CrossRef]
83. Ji, X.; Cao, Y.; Tai, Y.; Wang, C.; Li, J.; Huang, F. Real-world super-resolution via kernel estimation and noise injection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 466–467.
84. Bell-Kligler, S.; Shocher, A.; Irani, M. Blind super-resolution kernel estimation using an internal-gan. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019.
85. Luo, P.; Ren, J.; Peng, Z.; Zhang, R.; Li, J. Differentiable learning-to-normalize via switchable normalization. *arXiv* **2018**, arXiv:1806.10779.